On Properties of Locally Optimal Multiple Description Scalar Quantizers with Convex Cells

Sorina Dumitrescu and Xiaolin Wu
Department of Electrical and Computer Engineering
McMaster University, Hamilton, ON, Canada L8S 4K1
sorina/xwu@mail.ece.mcmaster.ca

Abstract- It is known that the generalized Lloyd method is applicable to locally optimal multiple description scalar quantizer (MDSQ) design. However, it remains unsettled when the resulting MDSQ is also globally optimal. We partially answer the above question by proving that for a fixed index assignment there is a unique locally optimal fixed-rate MDSQ of convex cells under Trushkin's sufficient conditions for the uniqueness of locally optimal fixed-rate single description scalar quantizer. This result holds for fixed-rate multiresolution scalar quantizer (MRSQ) of convex cells as well. Thus the well-known log-concave pdf condition can be extended to the multiple description and multiresolution cases.

Moreover we solve the difficult problem of optimal index assignment for fixed-rate MRSQ and symmetric MDSQ, when cell convexity is assumed. In both cases we prove that at optimality the number of cells in the central partition has to be maximal, as allowed by the side quantizer rates. As long as this condition is satisfied, any index assignment is optimal for MRSQ, while for symmetric MDSQ an optimal index assignment is proposed.

The condition of convex cells is also discussed. It is proved that cell convexity is asymptotically optimal for high resolution MRSQ, under the r^{th} power distortion measure.

Key words: convexity, index assignment, multiple descriptions, multiresolution, quantization.

Both authors were supported by Natural Sciences and Engineering Research Council of Canada. Parts of this work appeared in the Proceedings of 2005 IEEE Data Compression Conference , Snowbird, UT, March 2005, and the Proceedings of the 9th Canadian Workshop on Information Theory, Montreal, June 2005.

I. INTRODUCTION

The problem of multiple description coding (MDC) was first posed by Wyner at the 1979 IEEE Information Theory Workshop. Recent years have seen greatly intensified research efforts on MDC, which were primarily driven by the pressing needs of robust communications over lossy IP and wireless networks. A popular MDC technique is multiple description quantization.

A multiple description quantizer (MDQ) consists of a set of K encoders, also called *side encoders*, and a set of $2^K - 1$ decoders. Each encoder generates a different description, called *side description*. Each of the K side descriptions can be separately decoded to a reconstruction of a certain fidelity, by a so-called *side decoder*. Furthermore, any non-empty subset $\mathcal{I}$, $|\mathcal{I}| \geq 2$, of the K side descriptions can also be jointly decoded by a so-called *joint decoder*. The reconstruction improves as more side descriptions are received and they collaborate to refine the source. Each pair of a side encoder and the corresponding side decoder of the MDQ forms a quantizer (called *side quantizer*). Moreover, for each set $\mathcal{I}$, $|\mathcal{I}| \geq 2$, of descriptions there is an implicit encoder related to the side encoders for the descriptions in $\mathcal{I}$. This encoder together with the decoder for set $\mathcal{I}$ constitute a *joint quantizer*. The side and joint quantizers are called components of the MDQ. A component of particular importance is the *central quantizer*, which corresponds to the whole set of K descriptions.

This paper is concerned with the design and properties of optimal multiple description scalar quantizers. We will use the abbreviation K-DSQ for a K-description scalar quantizer. The performance of a K-DSQ is measured by the expected distortion of the reconstructed source at the receiver, where the expectation is taken over all possible sets of descriptions received. Thus, it is a weighted sum of the distortions of all decoders. The objective of optimal K-DSQ design is to minimize the expected distortion subject to rate constraints on the K side descriptions.

A multiresolution (progressively refinable) scalar quantizer (MRSQ) is a special case of K-DSQ, where a prefix condition has to be met. Namely, side description i can be decoded only jointly with *all* of the side descriptions 1, 2, through $i-1$, for any $1 \leq i \leq K$. In other words, the set of component decoders is restricted to the decoders associated to the sets of descriptions $\{1, 2, \cdots, i\}$ for all $1 \leq i \leq K$.

Another important case, which is commonly treated in the literature, is that of symmetric K-DSQ, when all K side descriptions have the same rate, and any two sets of descriptions of equal size have the same probability of being received.

A few techniques have been proposed for optimal K-DSQ design in both fixed-rate and entropy-constrained cases. The main design methods can be classified into two types: generalized Lloyd algorithms [26], [27], [4], [17], [12], and combinatorial algorithms [22], [23], [5], [6], [7], [8], [16]. The first type is a generalization of Lloyd's method [21] for fixed-rate scalar quantizer design. This method alternatively optimizes the decoder and the encoder, when the other component is fixed. Since the sequence of expected distortions is non-increasing, the algorithm eventually converges to a local optimum. The generalized Lloyd's method has been applied to the design of 2-DSQ with balanced descriptions (i.e., symmetric 2-DSQ in our terminology) [26], [27]. It was also used for the design of multiresolution scalar quantizers [4], [17] and multiresolution vector quantizers [9]. The design of more general K-DSQ by this approach was covered in [12].

The combinatorial methods address the design of optimal K-DSQ for discrete distributions, under the constraint of convex cells. These algorithms ensure globally optimal solution under the above convexity constraint. Specifically, in [22], [23] the problem is modeled as a shortest path problem in a weighted directed acyclic graph. In [7], [8], [6] a strong monotonicity of a general class of distortion functions is exploited to speed up the shortest path computation for fixed-rate 2-DSQ [7], [6] and symmetric fixed-rate K-DSQ [8]. Similar ideas are used in [5] to accelerate the dynamic programming in the optimal design of fixed-rate MRSQ. Unfortunately, the cell convexity may preclude optimal solutions [10]. Moreover, the fast matrix search algorithm for K-DSQ can still be too expensive if K is large (the complexity is exponential in K in general).

In contrast, the locally descent algorithms are easier to implement and have lower complexity. But they only converge to a locally optimal solution in general. However, this limitation becomes nonexistent if the underlying distortion function has a unique local minimum. The main question to be answered by this paper is under what conditions the local minimum is unique, and hence the generalized Lloyd method for K-DSQ design is globally optimal.

Sufficient conditions for the uniqueness of a local optimum were studied in the case of fixed-rate single description scalar quantization [11], [24], [18], [19], [25]. The most general sufficient conditions are the ones given by Trushkin in [24], which were shown to be satisfied if the source pdf is log-concave [25].

It is unknown up to now, however, whether a similar result holds for fixed-rate K-DSQ's. One of the contributions of this paper is to prove that Trushkin's conditions also ensure the uniqueness of locally optimal fixed-rate K-DSQ's with convex cells (convex K-DSQ's).

A central mechanism of K-DSQ is index assignment (IA) introduced in [26]. IA labels each central quantizer cell by an ordered K-tuple of indexes corresponding to the side quantizers cells whose intersection equals that central cell. Thus, the system of K side encoders is uniquely specified by the central partition (the partition induced by the central quantizer) and the IA. Our results on the uniqueness of locally optimal convex K-DSQ's hold with respect to fixed IA. The problem of optimal IA is notoriously difficult, and it is solved by us for two notable cases of K-DSQ with convex cells. Specifically, we present an optimal IA for fixed-rate convex symmetric K-DSQ. Moreover, for fixed-rate MRSQ of convex cells we show that any index assignment is optimal as long as the number of cells in the central partition is the maximum possible under the rate constraints. Furthermore, for both cases, we prove that the requirement for the number of cells in the central partition to be maximal is necessary at optimality.

As to cell convexity, it was shown [10] that there are discrete distributions and weighting schemes for different side and joint quantizers such that optimal K-DSQ has non-convex cells. It is interesting to know when the cell convexity does not precludes optimality. Qualitatively, optimal K-DSQ will necessarily have convex cells when the weights of side quantizers are large enough relative to the weights of the joint quantizers, i.e., the optimization emphasizes on the side quantizers rather than the joint ones. This intuition was validated in [6] for fixed-rate symmetric 2-DSQ's of high rates, and r-th power distortion measure. Another contribution of this paper is to show that the cell convexity of fixed-rate MRSQ does not preclude optimality for high rates and r-th power distortion measure, regardless the weighting scheme.

The next section introduces the definitions and notations used throughout the paper. Section 3 presents the necessary conditions for a locally optimal fixed-rate convex K-DSQ. Section 4 states and proves a key result of this paper: For convex and strictly increasing error functions, the sufficient conditions given by Trushkin [24] for the uniqueness of a locally optimal fixed-rate scalar quantizer are also sufficient for the uniqueness of a locally optimal fixed-rate convex K-DSQ, with respect to a given IA. Section 5 turns to the problem of optimal IA, in which we derive a necessary condition for optimality (within the class of fixed-rate convex MRSQ): the central partition of MRSQ has to have the largest possible number of cells allowed by the rate constraints. Further, any IA is optimal as long as this condition is satisfied. Next we prove in Section 6 that at optimality, the fixed-rate convex symmetric K-DSQ must also have the maximal number of cells in the central partition, and present an optimal IA. The proposed IA is a generalization of the staggered IA for two descriptions. Precisely, it requires that the j^{th} threshold of side partition i be the $\big((j-1)K+i\big)^{th}$ threshold in the central partition. Moreover, we show that if the error function is continuously differentiable, this IA is the unique optimal IA, up to a permutation of side quantizers. In Section 7 we discuss the cell convexity condition and show, based on the high-resolution analysis of optimal scalar quantization [2], [1], [3], [20], [14], that at high rates optimal fixed-rate MRSQ has convex cells for the r^{th} power distortion measure. Section 7 concludes the paper.

II. Definitions, Notations, Problem Formulation

Let X be a continuous random variable with probability density function (pdf, for short) $p(x)$. In this work we assume that the pdf $p(x)$ satisfies the following condition.

Condition A. There is an open interval (V, W), $-\infty \le V < W \le \infty$, such that $p(x)$ is continuous and positive inside this interval and $p(x) = 0$ outside this interval. Denote $\mathcal{A} = [V, W] \cap \mathbb{R}$.

We consider a distortion function $d(x, y) = f(|x - y|)$, where $f(\cdot)$ satisfies the condition stated below.

Condition B. $f(\cdot)$ is a nonnegative convex function with its only null point in 0. Consequently, $f(\cdot)$ is continuous and strictly increasing. Additionally, for any $y \in \mathbb{R}$ the following inequality holds

$$\int_V^W f(|y - x|)p(x)dx < +\infty.$$

A scalar quantizer Q of M cells is a partition of the alphabet set $\mathcal{A}$ into M non-empty sets $C_1, C_2, \cdots, C_M$, called cells, together with a set of representation values $y_1, \cdots, y_M \in \mathcal{A}$. The distortion of the quantizer is defined by

$$D(Q) = \sum_{i=1}^{M} \int_{C_i} d(x, y_i)p(x)dx = \sum_{i=1}^{M} \int_{C_i} f(|x - y_i|)p(x)dx. \tag{1}$$

The quantizer is also associated with a rate, denoted by $R(Q)$. In the case of fixed-rate quantizer, $R(Q) = \log_2 M$; in the case of entropy-constrained quantizer, $R(Q) = \sum_{i=1}^{M} P(C_i) \log_2 \frac{1}{P(C_i)}$, where $P(C_i) = \int_{C_i} p(x)dx$. The problem of optimal quantizer design is to minimize $D(Q)$ given a target quantizer rate $R(Q) = R_0$.

Now consider $K \geq 2$ different scalar quantizers $Q_1, Q_2, \cdots, Q_K$, called side quantizers, and define a K-description scalar quantizer (K-DSQ) $\mathbf{Q}$ to be a system of $2^K - 1$ scalar quantizers $Q_{\mathcal{I}}$, for $\mathcal{I} \subseteq \mathcal{K} = \{1, 2, \cdots, K\}$, $\mathcal{I} \neq \emptyset$, such that for each $\mathcal{I} = \{i_1, \cdots, i_s\}$ with $s \geq 2$, the partition of the alphabet $\mathcal{A}$ by quantizer $Q_{\mathcal{I}}$ equals the intersection of the partitions of quantizers $Q_{i_1}, \cdots, Q_{i_s}$. The quantizer $Q_{\mathcal{K}}$, which has the highest resolution among all joint quantizers, is called the central quantizer. As K-DSQ is a means of networked source coding to utilize channel diversity, it is natural to define the expected distortion $\bar{D}(\mathbf{Q})$ of the K-DSQ $\mathbf{Q}$ to be

$$\bar{D}(\mathbf{Q}) = \sum_{\mathcal{I} \subseteq \mathcal{K}, \mathcal{I} \neq \emptyset} \omega_{\mathcal{I}} D(Q_{\mathcal{I}}), \tag{2}$$

where each component quantizer $Q_{\mathcal{I}}$ is assigned a weight $\omega_{\mathcal{I}} \geq 0$. Typically, in practice the weight $\omega_{\mathcal{I}}$ has the meaning of the probability that only the subset of side descriptions $\mathcal{I}$ is available for source reconstruction. Note that in (2) the term for no descriptions is omitted since it does not affect the optimal design of K-DSQ. For convenience of further formulations we also define $\omega_{\emptyset} = 0$.

The K-DSQ is said to be fixed-rate/entropy-constrained if all side quantizers are fixed-rate/entropy-constrained.

The problem of optimal fixed-rate/entropy-constrained K-DSQ design is to minimize the expected distortion (2) over all possible K side quantizers, given the weights $\omega_{\mathcal{I}}$, and given the target rates $R(Q_i) = R_i$, $1 \leq i \leq K$, of the side quantizers. Note that in the case of fixed-rate K-DSQ, the constraints on the rates are equivalent to imposing to each side quantizer Q_i to have $M_i = 2^{R_i}$ cells.

Note that only the quantizers $Q_{\mathcal{I}}$ with $\omega_{\mathcal{I}} \neq 0$, contribute to the expected distortion. Therefore we call them *active* components of the K-DSQ. We will require that the central quantizer of a K-DSQ be an active component, i.e., $\omega_{\mathcal{K}} > 0$.

A K-DSQ is called symmetric if $R_1 = R_2 = \cdots = R_K$ and $\omega_{\mathcal{I}} = \omega_{\mathcal{I}'}$ for all $\mathcal{I}, \mathcal{I}' \subseteq \mathcal{K}$, such that $|\mathcal{I}| = |\mathcal{I}'|$. We also require that the side quantizers of symmetric K-DSQ are active components, i.e., $\omega_{\{1\}} > 0$.

The above definition of K-DSQ also includes multiresolution scalar quantizers (MRSQ). Precisely, an MRSQ of K refinement stages is a K-DSQ whose active components are $Q_1, Q_{\{1,2\}}, \cdots, Q_{\{1,\cdots,i\}}, \cdots, Q_{\mathcal{K}}$.

A cell is said to be convex if it is a convex set, i.e., an interval of the real line. A scalar quantizer is called convex, or regular as referred in some literature, if all of its cells are convex. A K-DSQ is said to

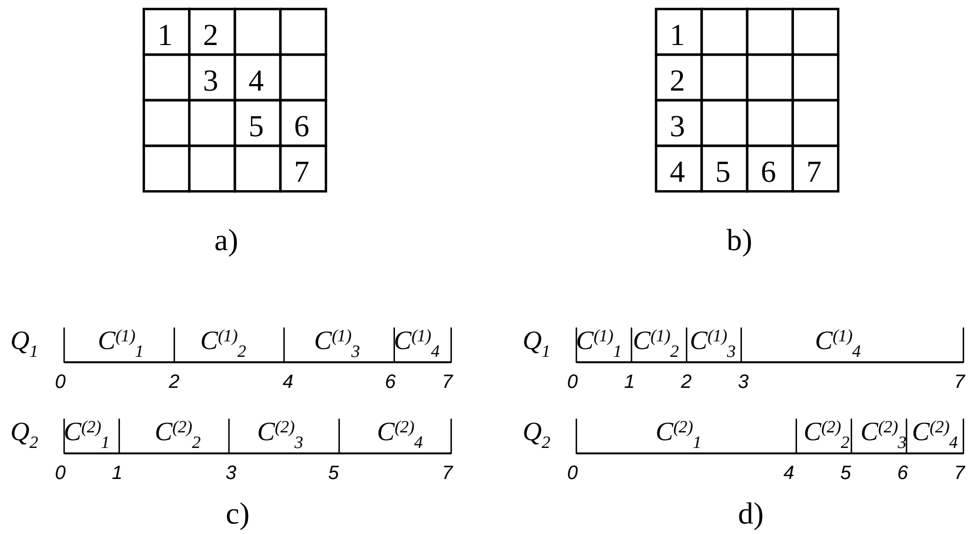

Fig. 1. Two different IA's for convex 2DSQ and their corresponding side quantizers partitions: c) partitions for IA of a); d) partitions for IA of b).

be convex if all its active quantizers are convex. Note that in a convex K-DSQ, not all side quantizers are necessarily convex, but only if they are active components. For example, in a convex MRSQ only the side quantizer Q_1 is active, therefore all the others may have non-convex cells.

Assume that the K-DSQ has convex cells in the central partition and denote them by $C_1, C_2, \cdots, C_M$, the indexing being consistent with their order from left to right. Further, for each $i, 1 \leq i \leq K$, let M_i denote the number of cells of the side quantizer Q_i and let $C_1^{(i)}, C_2^{(i)}, \cdots, C_{M_i}^{(i)}$, denote its cells. The index assignment of the K-DSQ is the mapping $h : \{1, \cdots, M\} \rightarrow \{1, \cdots, M_1\} \times \cdots \times \{1, \cdots, M_K\}$, such that $h(l) = (j_1, j_2, \cdots, j_K)$ if and only if $C_l = C_{j_1}^{(1)} \cap C_{j_2}^{(2)} \cap \cdots \cap C_{j_K}^{(K)}$. In other words, the IA is the function which assigns to each l the K-tuple of indices of side cells whose intersection equals the l^{th} central cell. Note that the the central partition together with the IA uniquely determine the partitions of all component quantizers. This is because for any j and i the side cell $C_j^{(i)}$ has to be the union of all C_l for which the ith component of $h(l)$ equals j.

It is true that in the case of convex K-DSQ the convexity requirement already imposes constraints on the IA. However, the variety of eligible IA's may still be large enough to make an exhaustive search for the optimal IA unattractive.

Next we illustrate the relevance of IA for convex K-DSQ's by considering two examples, one for fixed-rate symmetric 2-DSQ and one for fixed-rate MRSQ with 2 refinement stages. The number of cells in the central partition is $M = 7$ in the first example and $M = 6$ in the second one. In both cases we assume a uniform distribution over $[V, W] = [0, M]$, and the central partition to be uniform, consequently, $C_l = (l-1, l]$ for $2 \leq l \leq M$, and $C_1 = [0, 1]$. In these examples we will consider the squared error as distortion measure and the midpoint of each cell as its representation value.

We will represent graphically each IA as a table with M_1 rows and M_2 columns, where some positions are filled with the integers $l, 1 \leq l \leq M$, while others are empty. Precisely, integer l is placed on row i and column j if and only if $h(l) = (i, j)$. This table will be referred to as the IA matrix as in [26].

Example 1. Relevance of IA for Convex Symmetric 2-DSQ. Consider the two IA's depicted in Figure 1 a) and b), where $M_1 = M_2 = 4$. Each of the IA's induces a convex 2-DSQ since the cells of both side quantizers are convex. The side partitions corresponding to each of the two IA's are represented in Figure 1 c) and d), respectively. The side distortions for case c) are $D(Q_1) = D(Q_2) = \frac{25}{84}$, while for case d) are $D(Q_1) = D(Q_2) = \frac{67}{84}$. Since the central distortion is the same in both cases, it follows that the 2-DSQ of case c) has smaller expected distortion.

Example 2. Relevance of IA for Convex 2-stage MRSQ. Consider the three IA's illustrated in Figure 2 a)-c). Here $M_1 = 2$ and $M_2 = 4$. Each of these IA's defines a 2-stage MRSQ where Q_2 has non-convex cells, while Q_1 has only convex cells. Since only Q_1 and $Q_{\{1,2\}}$ must have convex cells for the MRSQ to qualify as convex, it follows that all three IAs correspond to convex MRSQ's. Their side partitions are depicted in Figure 2 d)-f), respectively. Note that the side distortion $D(Q_1)$ is smaller in case d) than in the other two cases since, as it is well known, for the uniform distribution, the uniform quantizer is strictly better than a non-uniform one. Further, because the side quantizer Q_2 does not affect the overall expected distortion, while the central quantizer is the same in all three situations, it follows that the MRSQ of case d) has the best performance. On the other hand, note that the MRSQ's of cases e) and f) have different quantizers Q_2, but identical quantizers Q_1 and $Q_{\{1,2\}}$, in other words their active quantizers coincide. It follows that their performance is the same.

III. Necessary Conditions for a Locally Optimal K-DSQ

In this paper the scope of our inquiry is confined to the class of fixed-rate convex K-DSQ's. For succinctness of presentation we will drop the qualifier "fixed-rate", consequently only use the term

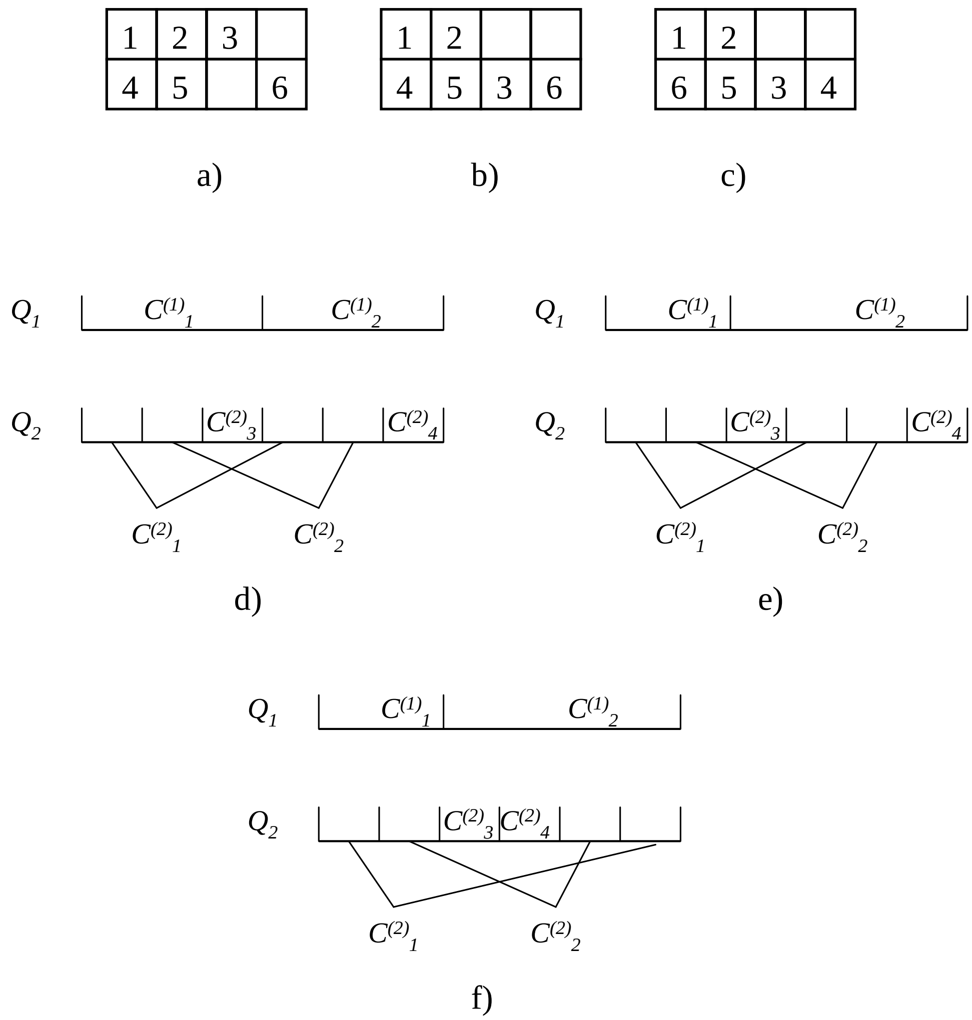

Fig. 2. Three different IA's for convex MRSQ and their corresponding side quantizers partitions: d) partition for IA of a); e) partition for IA of b); f) partition for IA of c).

convex K-DSQ in sequel.

Design algorithms generalized from Lloyd's method [21], starting from a configuration of the K-DSQ, alternatively fix the encoder to optimize the decoder and fix the decoder to optimize the encoder. Since at each iteration the expected distortion does not increase, such an algorithm eventually converges to a locally optimal convex K-DSQ. We next present the necessary conditions for a locally optimal convex K-DSQ.

Let $\mathbf{Q}$ be a convex K-DSQ. Denote by $M_{\mathcal{I}}$ the number of cells of quantizer $Q_{\mathcal{I}}$, for $\mathcal{I} \subseteq \mathcal{K}$, $\mathcal{I} \neq \emptyset$ and write $M = M_{\mathcal{K}}$. Let $(u_{i-1}, u_i]$, for $1 \leq i \leq M$, be the cells of the central quantizer, where

$V = u_0 < u_1 < u_2 < \cdots < u_M = W$. The values $u_i, 1 \le i \le M-1$, are called thresholds. For each active quantizer $Q_\mathcal{I}$ there are indices $j_{\mathcal{I},0} = 0 < j_{\mathcal{I},1} < j_{\mathcal{I},2} < \cdots < j_{\mathcal{I},M_\mathcal{I}} = M$, such that the cells of $Q_\mathcal{I}$ are $(u_{j_{\mathcal{I},k-1}}, u_{j_{\mathcal{I},k}}]$, for $1 \le k \le M_\mathcal{I}$. Denote by $y_{\mathcal{I},k}$ the reproduction value corresponding to cell $(u_{j_{\mathcal{I},k-1}}, u_{j_{\mathcal{I},k}}]$ of $Q_\mathcal{I}$. Then (1) and (2) imply that

$$\bar{D}(\mathbf{Q}) = \sum_{\mathcal{I}\subseteq\mathcal{K},\omega_\mathcal{I}\neq 0} \omega_\mathcal{I} \sum_{k=1}^{M_\mathcal{I}} \int_{u_{j_{\mathcal{I},k-1}}}^{u_{j_{\mathcal{I},k}}} f(|x - y_{\mathcal{I},k}|)p(x)dx. \tag{3}$$

Optimum decoder condition. When the encoder is fixed, the thresholds u_i and the indeces $j_{\mathcal{I},k}$ are fixed. Thus, the decoder is optimum if and only if the following is satisfied

$$\int_{u_{j_{\mathcal{I},k-1}}}^{u_{j_{\mathcal{I},k}}} f(|x - y_{\mathcal{I},k}|)p(x)dx = \min_{y\in\mathcal{A}} \int_{u_{j_{\mathcal{I},k-1}}}^{u_{j_{\mathcal{I},k}}} f(|x-y|)p(x)dx,$$

for all $\mathcal{I}$ such that $\omega_\mathcal{I} > 0$, and $k, 1 \le k \le M_\mathcal{I}$.

As shown by Trushkin [24], for every $V \le a < b \le W$, the function $D_{a,b}(y) = \int_a^b f(|x-y|)p(x)dx$, defined for every $y \in [V, W]$, achieves its minimum in some unique point $\mu(a, b)$, situated inside the interval (a, b). This value is called generalized centroid. Consequently, the optimum decoder condition is that

$$y_{\mathcal{I},k} = \mu(u_{j_{\mathcal{I},k-1}}, u_{j_{\mathcal{I},k}}),\ \text{for all } \mathcal{I},\ \omega_\mathcal{I} > 0, \text{ and } k, 1 \le k \le M_\mathcal{I}. \tag{4}$$

Note that for the case of squared error distortion measure, $\mu(a, b)$ is the conditional mean of the interval (a, b), i.e.,

$$\mu(a, b) = \frac{\int_a^b xp(x)dx}{\int_a^b p(x)dx}. \tag{5}$$

Optimum encoder condition. Here we derive a necessary condition for optimal encoder, given fixed decoder and IA.

Define a function $G(\mathbf{u})$ on the set $\mathcal{O}_M$ of $(M-1)$-dimensional vectors $\mathbf{u} = (u_1, u_2, \cdots, u_{M-1})$ satisfying $V < u_1 < u_2 < \cdots < u_{M-1} < W$:

$$G(\mathbf{u}) = \sum_{\mathcal{I}\subseteq\mathcal{K},\omega_\mathcal{I}\neq 0} \omega_\mathcal{I} \sum_{k=1}^{M_\mathcal{I}} \int_{u_{j_{\mathcal{I},k-1}}}^{u_{j_{\mathcal{I},k}}} f(|x - y_{\mathcal{I},k}|)p(x)dx.$$

The encoder is optimal given the decoder and the index assignment if and only if $G(\cdot)$ takes its minimum value over $\mathcal{O}_M$ at $\mathbf{u}$. Since $G(\cdot)$ is continuous and differentiable and $\mathcal{O}_M$ is an open set, a necessary condition for $G(\mathbf{u})$ to be the minimum over $\mathcal{O}_M$ is:

$$\frac{\partial G}{\partial u_i}(\mathbf{u}) = 0,\ \text{for any } i, 1 \le i \le M-1. \tag{6}$$

Fix an arbitrary $i, 1 \leq i \leq M-1$. Denote by $\mathcal{S}_i$ the set of subsets of indices $\mathcal{I} \subseteq \mathcal{K}$, such that $\omega_{\mathcal{I}} \neq 0$ and u_i is a threshold of quantizer $Q_{\mathcal{I}}$. Thus, for each $\mathcal{I} \in \mathcal{S}_i$ there is an integer $k(\mathcal{I}, i)$ such that $j_{\mathcal{I},k(\mathcal{I},i)} = i$ (i.e., u_i is the $k(\mathcal{I}, i)^{th}$ threshold of quantizer $Q_{\mathcal{I}}$). Then

$$G(\mathbf{u}) = T + \sum_{\mathcal{I} \in \mathcal{S}_i} \omega_{\mathcal{I}} \left(\int_{u_{j_{\mathcal{I},k(\mathcal{I},i)-1}}}^{u_i} f(|x - y_{\mathcal{I},k(\mathcal{I},i)}|)p(x)dx + \int_{u_i}^{u_{j_{\mathcal{I},k(\mathcal{I},i)+1}}} f(|x - y_{\mathcal{I},k(\mathcal{I},i)+1}|)p(x)dx \right),$$

where the term T does not depend on u_i. It follows that

$$\frac{\partial G}{\partial u_i}(\mathbf{u}) = \sum_{\mathcal{I} \in \mathcal{S}_i} \omega_{\mathcal{I}} p(u_i) \left(f(|u_i - y_{\mathcal{I},k(\mathcal{I},i)}|) - f(|u_i - y_{\mathcal{I},k(\mathcal{I},i)+1}|) \right).$$

Since $p(u_i) \neq 0$, the necessary condition for optimal encoder (6) becomes

$$\sum_{\mathcal{I} \in \mathcal{S}_i} \omega_{\mathcal{I}} f(|u_i - y_{\mathcal{I},k(\mathcal{I},i)}|) = \sum_{\mathcal{I} \in \mathcal{S}_i} \omega_{\mathcal{I}} f(|u_i - y_{\mathcal{I},k(\mathcal{I},i)+1}|) \text{ for any } i, 1 \leq i \leq M-1. \tag{7}$$

The K-DSQ obtained at convergence must simultaneously satisfy (4) and (7). By combining these two conditions we obtain

$$\sum_{\mathcal{I} \in \mathcal{S}_i} \omega_{\mathcal{I}} f(|u_i - \mu(u_{j_{\mathcal{I},k(\mathcal{I},i)-1}}, u_i)|) = \sum_{\mathcal{I} \in \mathcal{S}_i} \omega_{\mathcal{I}} f(|u_i - \mu(u_i, u_{j_{\mathcal{I},k(\mathcal{I},i)+1}})|), 1 \leq i \leq M-1, \tag{8}$$

which is the necessary condition for local optimum with respect to the IA. From now on we simply refer to it as the necessary condition for local optimum, being understood that this condition takes different forms for different IA's.

IV. Sufficient Conditions for Uniqueness of a Locally Optimal convex K-DSQ

Sufficient conditions for the global optimality of a locally optimal fixed-rate scalar quantizer were investigated in [11], [24], [18], [19], [25]. The sufficient conditions found by Fleischer [11] in the case of squared error distortion, require that $p(x)$ be differentiable and the derivative of $\log_2 p(x)$ be strictly decreasing. Trushkin considered a more general distortion measure, namely $d(x, y) = g(x, |x - y|)$, such that $g(x, \eta)$ is convex in η, has a unique zero point for each x and is continuous in x [24]. He formulated sufficient conditions for the uniqueness of a locally optimal quantizer which do not require differentiability of $p(x)$. The log-concavity of $p(x)$ satisfies these conditions when $g(x, \eta) = \phi(x)\eta^2$ or $g(x, \eta) = \phi(x)|\eta|$. Kieffer [19] proved the sufficiency of log-concavity for the family of distortion functions $d(x, y) = f(|x - y|)$, where $f(\cdot)$ is increasing, convex and continuously differentiable. Finally, Trushkin [25] extended this result by showing that the requirement of continuous differentiability for

$f(\cdot)$ can be dropped. To our best knowledge the conditions formulated by Trushkin in [24] are the most general sufficient conditions for uniqueness of local optimal scalar quantizer so far.

As specified in Section 2 our distortion function is $d(x, y) = f(|x - y|)$, where $f(\cdot)$ is a nonnegative convex function with its only null point in 0, i.e., $f(\cdot)$ is continuous and strictly increasing. The next theorem states that under the same conditions for $p(x)$ as in [24, Theorem 1], there is at most one locally optimal K-DSQ of a given IA. An immediate consequence is that the log-concavity of $p(x)$ suffices to ensure this uniqueness.

Theorem 1. Assume that the pdf $p(\cdot)$ satisfies Condition A and the error function $f(\cdot)$ satisfies Condition B. Additionally assume that conditions $T1 - T4$ stated below hold. Then there is at most one $\mathbf{u} \in \mathcal{O}_M$ which satisfies (8).

$T1$) For any $V < x_0 < x_1 < W$, $x_0 - \mu(V, x_0) \leq x_1 - \mu(V, x_1)$.

$T2$) For any $V < x_0 < x_1 < W$, $\mu(x_0, W) - x_0 \geq \mu(x_1, W) - x_1$.

$T3$) For any $V < x_0 < z_0 < W$, $V < x_1 < z_1 < W$, such that $x_0 \leq x_1$, there is

$$\mu(x_0, z_0) - x_0 \leq \mu(x_1, z_1) - x_1 \Rightarrow z_0 - \mu(x_0, z_0) \leq z_1 - \mu(x_1, z_1). \tag{9}$$

Moreover, if the left inequality is strict, then so is the right one.

$T4$) For any positive integer m and any two sets of values $V < x_1 < \cdots < x_m < W$, $V < z_1 < \cdots < z_m < W$, such that $x_i < z_i$, $1 \leq i \leq m$, and $\mu(x_i, x_{i+1}) - x_i \leq \mu(z_i, z_{i+1}) - z_i$, $1 \leq i \leq m - 1$, at least one of the following inequalities holds:

$T4.1$) $x_1 - \mu(V, x_1) < z_1 - \mu(V, z_1)$;

$T4.2$) $\mu(x_m, W) - x_m > \mu(z_m, W) - z_m$;

$T4.3$) for some $i, 1 \leq i \leq m - 1$, $x_{i+1} - \mu(x_i, x_{i+1}) < z_{i+1} - \mu(z_i, z_{i+1})$.

In order to prove Theorem 1 we first write the necessary condition for local optimum (8) in a simpler form. Denote respectively by $\gamma_i^{(1)}(\mathbf{u})$ and $\gamma_i^{(2)}(\mathbf{u})$, the expression in the left and right hand side of equation (8), which is rewritten as

$$\gamma_i^{(1)}(\mathbf{u}) = \gamma_i^{(2)}(\mathbf{u}) \text{ for all } i, 1 \leq i \leq M - 1. \tag{10}$$

There are coefficients $\alpha_{i,j} \geq 0$, for $0 \leq j < i$, $\alpha_{i,i-1} > 0$ (because $\omega_{\mathcal{K}} > 0$), and coefficients $\beta_{i,l} \geq 0$, for $i < l \leq K$, $\beta_{i,i+1} > 0$ (because $\omega_{\mathcal{K}} > 0$), such that

$$\gamma_i^{(1)}(\mathbf{u}) = \sum_{j=0}^{i-1} \alpha_{i,j} f(u_i - \mu(u_j, u_i)), \text{ and } \gamma_i^{(2)}(\mathbf{u}) = \sum_{l=i+1}^{M} \beta_{i,l} f(\mu(u_i, u_l) - u_i). \tag{11}$$

Note that a single description quantizer can be considered as a special case of K-DSQ with $K = 1$. Then the central partition coincides with the side partition 1 and $M = M_1$. Consequently, conditions (10) also characterize a locally optimal convex scalar quantizer. However, the corresponding functions $\gamma_i^{(1)}(\cdot)$ and $\gamma_i^{(2)}(\cdot)$ are simpler. Each of them is only a function of two consecutive thresholds and the summations in (11) contain a single term. Trushkin's approach to prove the uniqueness of a locally optimal quantizer under the conditions of Theorem 1, was to show that, if two locally optimal points $\mathbf{u}$ and $\mathbf{u}'$ have the first $k-1$ components equal and $u_k < u'_k$, then the inequality "$<$" propagates to the rest of the components, i.e., $u_i < u'_i$ holds for all $i > k$. The proof was completed by showing that relation $u_{M-1} < u'_{M-1}$ leads to a contradiction. In the case of K-DSQ, the summations in (11) contain more than one term and $\gamma_i^{(1)}(\cdot)$ and $\gamma_i^{(2)}(\cdot)$ are functions of more thresholds, facts which make the problem more complex. In this case if two locally optimal points $\mathbf{u}$ and $\mathbf{u}'$ have the first $k-1$ components equal and $u_k < u'_k$, then inequalities $u_i < u'_i$ do not necessarily hold for all $i > k$. Moreover, such inequalities would not be sufficient to reach a contradiction. Our approach is to show that a more complex condition is propagated to some values of $i > k$ until a contradiction is reached. To proceed with the proof we first present two lemmas.

Lemma 1. If Conditions A, B and $T3$ hold, then for any $V < x_0 < z_0 < W$, $V < x_1 < z_1 < W$, such that $z_0 < z_1$, there is

$$z_0 - x_0 \leq z_1 - x_1 \Rightarrow z_0 - \mu(x_0, z_0) \leq z_1 - \mu(x_1, z_1) \tag{12}$$

Moreover, if the first inequality is strict, so is the second one.

Proof. Note first that implication (9) of $T3$ is equivalent to

$$x_1 - x_0 \leq \mu(x_1, z_1) - \mu(x_0, z_0) \Rightarrow \mu(x_1, z_1) - \mu(x_0, z_0) \leq z_1 - z_0. \tag{13}$$

Likewise, relation (12) is equivalent to

$$x_1 - x_0 \leq z_1 - z_0 \Rightarrow \mu(x_1, z_1) - \mu(x_0, z_0) \leq z_1 - z_0. \tag{14}$$

Assume now that the hypothesis of Lemma 1 holds. Next we need to distinguish between two cases.

Case a) $x_0 \leq x_1$. If $x_1 - x_0 \leq \mu(x_1, z_1) - \mu(x_0, z_0)$, then the second inequality in (14) follows by $T3$ (according to (13)). If $x_1 - x_0 > \mu(x_1, z_1) - \mu(x_0, z_0)$, then combining this relation with $x_1 - x_0 \leq z_1 - z_0$, again $\mu(x_1, z_1) - \mu(x_0, z_0) \leq z_1 - z_0$ follows. Now let us assume that $x_1 - x_0 < z_1 - z_0$ and that $\mu(x_1, z_1) - \mu(x_0, z_0) = z_1 - z_0$. Then the first inequality in (13) holds and it is strict, and by $T3$ the second one is strict too, thus leading to a contradiction. Hence, the claim of Lemma 1 is proved.

Case b) $x_1 < x_0$. Because the function $\mu(\cdot, \cdot)$ is non-decreasing in both variables [24], $x_1 < x_0$ and $z_0 < z_1$, imply that $\mu(x_0, z_0) \geq \mu(x_1, z_0)$ and $\mu(x_1, z_0) \leq \mu(x_1, z_1)$. The last relation can be rewritten as $0 = x_1 - x_1 \leq \mu(x_1, z_1) - \mu(x_1, z_0)$ and by $T3$ (via (13)) it implies that $\mu(x_1, z_1) - \mu(x_1, z_0) \leq z_1 - z_0$, and that the equality holds only if $\mu(x_1, z_1) - \mu(x_1, z_0) = 0$. On the other hand, equality in the last two relations cannot be reached simultaneously because $z_1 - z_0 > 0$. Consequently, we have $\mu(x_1, z_1) - \mu(x_1, z_0) < z_1 - z_0$. Using further the fact that $\mu(x_1, z_0) - \mu(x_1, z_1) \leq 0$, we obtain

$$\mu(x_1, z_1) - \mu(x_0, z_0) = \big(\mu(x_1, z_1) - \mu(x_1, z_0)\big) + \big(\mu(x_1, z_0) - \mu(x_0, z_0)\big) < z_1 - z_0,$$

which concludes the proof. $\square$

In order to state the next lemma we introduce a definition first. Consider two arbitrary points $\mathbf{u} = (u_1, \cdots, u_{M-1})$, $\mathbf{u}' = (u'_1, \cdots, u'_{M-1}) \in \mathcal{O}_M$, and an integer $i, 1 \leq i \leq M-1$. We say that condition $C(i)$ is satisfied if and only if $u_i < u'_i$ and the following inequalities hold

$$u_i - u_j \leq u'_i - u'_j \text{ for all } j, 1 \leq j < i. \tag{15}$$

Note that condition $C(1)$ is the condition that $u_1 < u'_1$.

Lemma 2. Assume that Conditions A, B, $T1$, $T2$ and $T3$ hold. Let $\mathbf{u}$,$\mathbf{u}'$ be arbitrary points in $\mathcal{O}_M$, and let i be an integer $1 \leq i \leq M-2$. If condition $C(i)$ is satisfied and $\gamma_i^{(1)}(\mathbf{u}) = \gamma_i^{(2)}(\mathbf{u})$ and $\gamma_i^{(1)}(\mathbf{u}') = \gamma_i^{(2)}(\mathbf{u}')$, then at least one of the following holds.

$L1$) $\mu(u_i, u_{i+1}) - u_i = \mu(u'_i, u'_{i+1}) - u'_i$ and $C(i+1)$ is satisfied;

$L2$) there is some $k, 1 \leq k \leq M-1-i$, such that condition $C(i+k)$ is satisfied and $\gamma_{i+k}^{(1)}(\mathbf{u}) < \gamma_{i+k}^{(1)}(\mathbf{u}')$.

If in addition to condition $C(i)$, we have $\gamma_i^{(1)}(\mathbf{u}) < \gamma_i^{(1)}(\mathbf{u}')$, then necessarily $L2$ holds. (Note that $L1$ and $L2$ do not necessarily exclude each other.)

Proof. Because $C(i)$ is satisfied, it follows that $u_i < u'_i$ and inequalities (15) hold. Thus, by applying $T1$ for $j = 0$ and Lemma 1 for $j > 0$, we obtain that

$$u_i - \mu(u_j, u_i) \leq u'_i - \mu(u'_j, u'_i),$$

for all $j, 0 \leq j < i$. Because the function $f(\cdot)$ is strictly increasing, it follows that

$$f(u_i - \mu(u_j, u_i)) \leq f(u'_i - \mu(u'_j, u'_i)), \tag{16}$$

for all $j, 0 \leq j < i$. Using the expression (11) for $\gamma_i^{(1)}(\cdot)$ and the fact that all coefficients $\alpha_{i,j}$ are non-negative, we have $\gamma_i^{(1)}(\mathbf{u}) \leq \gamma_i^{(1)}(\mathbf{u}')$. This and the hypothesis $\gamma_i^{(1)}(\mathbf{u}) = \gamma_i^{(2)}(\mathbf{u})$ and $\gamma_i^{(1)}(\mathbf{u}') = \gamma_i^{(2)}(\mathbf{u}')$ further imply

$$\gamma_i^{(2)}(\mathbf{u}) \leq \gamma_i^{(2)}(\mathbf{u}'), \tag{17}$$

which is equivalent to

$$\sum_{l=i+1}^{M} \beta_{i,l} f(\mu(u_i, u_l) - u_i) \leq \sum_{l=i+1}^{M} \beta_{i,l} f(\mu(u'_i, u'_l) - u'_i). \tag{18}$$

Because $\beta_{i,l} \geq 0$ for all l, and $\beta_{i,i+1} > 0$, at least one of the following conditions must hold:

$S1$) $f(\mu(u_i, u_{i+1}) - u_i) = f(\mu(u'_i, u'_{i+1}) - u'_i)$;

$S2$) there is some $k, 1 \leq k \leq M - 1 - i$ such that $f(\mu(u_i, u_{i+k}) - u_i) < f(\mu(u'_i, u'_{i+k}) - u'_i)$.

Indeed, if $S1$ does not hold then either $f(\mu(u_i, u_{i+1}) - u_i) < f(\mu(u'_i, u'_{i+1}) - u'_i)$ is true, in which case $S2$ holds for $k = 1$, or $f(\mu(u_i, u_{i+1}) - u_i) > f(\mu(u'_i, u'_{i+1}) - u'_i)$ is valid, in which case $S2$ must hold for some $k, 1 < k \leq M - i$ because otherwise the inequality (18) would not be satisfied. But $k \neq M - i$ because from $u_i < u'_i$ and $T2$ it follows that $\mu(u_i, u_M) - u_i \geq \mu(u'_i, u'_M) - u'_i$ (recall that $u_M = u'_M = W$), which implies that $f(\mu(u_i, u_M) - u_i) \geq f(\mu(u'_i, u'_M) - u'_i)$.

If $S1$ holds, we have $\mu(u_i, u_{i+1}) - u_i = \mu(u'_i, u'_{i+1}) - u'_i$ from the strict monotonicity of $f(\cdot)$. Since $u_i < u'_i$ we can apply $T3$, and obtain $u_{i+1} - \mu(u_i, u_{i+1}) \leq u'_{i+1} - \mu(u'_i, u'_{i+1})$, and further $(u_{i+1} - \mu(u_i, u_{i+1})) + (\mu(u_i, u_{i+1}) - u_i) \leq (u'_{i+1} - \mu(u'_i, u'_{i+1})) + (\mu(u'_i, u'_{i+1}) - u'_i)$, hence $u_{i+1} - u_i \leq u'_{i+1} - u'_i$. Additionally, since inequalities (15) hold, it follows that $(u_{i+1} - u_i) + (u_i - u_j) \leq (u'_{i+1} - u'_i) + (u'_i - u'_j)$ for all $j, 1 \leq j < i$. Consequently, $u_{i+1} - u_j \leq u'_{i+1} - u'_j$ for all $j, 1 \leq j < i + 1$.

Also, the inequalities $u_{i+1} - u_i \leq u'_{i+1} - u'_i$ and $u_i < u'_i$ imply that $u_{i+1} < u'_{i+1}$. Thus, condition $C(i + 1)$ is satisfied and conclusion $L1$ follows.

If $S2$ holds, then the inequality $f(\mu(u_i, u_{i+k}) - u_i) < f(\mu(u'_i, u'_{i+k}) - u'_i)$ implies that $\mu(u_i, u_{i+k}) - u_i < \mu(u'_i, u'_{i+k}) - u'_i$. By applying $T3$ (we are allowed because $u_i < u'_i$), we obtain $u_{i+k} - \mu(u_i, u_{i+k}) < u'_{i+k} - \mu(u'_i, u'_{i+k})$. These two inequalities lead to $u_{i+k} - u_i < u'_{i+k} - u'_i$. Let k_0 denote the smallest $k \geq 1$ which satisfies the previous inequality. Since $u_i < u'_i$, it follows that $u_{i+k_0} < u'_{i+k_0}$, too. By the definition of k_0, for any $k', 0 \leq k' < k_0$, we have $u_{i+k'} - u_i \geq u'_{i+k'} - u'_i$. Corroborating with $u_{i+k_0} - u_i < u'_{i+k_0} - u'_i$, it follows that $(u_{i+k_0} - u_i) - (u_{i+k'} - u_i) < (u'_{i+k_0} - u'_i) - (u'_{i+k'} - u'_i)$, and hence $u_{i+k_0} - u_{i+k'} < u'_{i+k_0} - u'_{i+k'}$ for all $k', 0 \leq k' < k_0$.

On the other hand, for any $j, 1 \leq j < i$, since $u_i - u_j \leq u'_i - u'_j$ by (15), and $u_{i+k_0} - u_i < u'_{i+k_0} - u'_i$, it follows that $u_{i+k_0} - u_j < u'_{i+k_0} - u'_j$. Consequently, condition $C(i+k_0)$ is satisfied with strict inequalities:

$$u_{i+k_0} - u_j < u'_{i+k_0} - u'_j \text{ for all } j, 1 \leq j < i + k_0. \tag{19}$$

Since the above inequalities are strict, it follows from Lemma 1 together with the strict monotonicity of $f(\cdot)$, that $f(u_{i+k_0} - \mu(u_j, u_{i+k_0})) < f(u'_{i+k_0} - \mu(u'_j, u'_{i+k_0}))$ for all $j, 1 \leq j < i + k_0$. Moreover, T1 and the monotonicity of $f(\cdot)$ imply that $f(u_{i+k_0} - \mu(u_j, u_{i+k_0})) \leq f(u'_{i+k_0} - \mu(u'_j, u'_{i+k_0}))$ for $j = 0$ (note that $u_0 = u'_0 = V$). Since $\alpha_{i+k_0,j} \geq 0$, $0 \leq j < i + k_0$, and $\alpha_{i+k_0,i+k_0-1} > 0$, we obtain further that $\gamma^{(1)}_{i+k_0}(\mathbf{u}) < \gamma^{(1)}_{i+k_0}(\mathbf{u}')$. Thus, $L2$ follows.

If $\gamma^{(1)}_i(\mathbf{u}) < \gamma^{(1)}_i(\mathbf{u}')$, then inequality (17) has to be strict, hence (18) has to be strict, too. Then clearly, $S1$ cannot hold, hence $S2$ has to hold, and $L2$ follows. $\square$

Proof of Theorem 1.

Assume that there are two different points $\mathbf{u} = (u_1, \cdots u_{M-1}) \in \mathcal{O}_M$, and $\mathbf{u}' = (u'_1, \cdots, u'_{M-1}) \in \mathcal{O}_M$, for which (8) (or equivalently, (10)) holds. We show that this assumption leads to a contradiction.

Since $\mathbf{u} \neq \mathbf{u}'$, it follows that there is some $i, 1 \leq i \leq M-1$, such that $u_i \neq u'_i$. Let i_0 be the smallest i with this property. We assume without loss of generality that $u_{i_0} < u'_{i_0}$. Then clearly $C(i_0)$ is satisfied. Thus Lemma 2 can be applied. Moreover, according to $T1$ and Lemma 1, we obtain that $u_{i_0} - \mu(u_j, u_{i_0}) \leq u'_{i_0} - \mu(u'_j, u'_{i_0})$, which further implies by (16) that $f(u_{i_0} - \mu(u_j, u_{i_0})) \leq f(u'_{i_0} - \mu(u'_j, u'_{i_0}))$, for all $j, 0 \leq j < i_0$. Because the coefficients $\alpha_{i_0,j}$ are nonnegative it follows further that $\gamma^{(1)}_{i_0}(\mathbf{u}) \leq \gamma^{(1)}_{i_0}(\mathbf{u}')$. We distinguish further two cases: when $i_0 \geq 2$, and when $i_0 = 1$.

Case 1. $i_0 \geq 2$. Because $V < u_{i_0-1} = u'_{i_0-1} < u_{i_0} < u'_{i_0}$, it follows that $u_{i_0} - u_{i_0-1} < u'_{i_0} - u'_{i_0-1}$.

Lemma 1 implies further that $u_{i_0} - \mu(u_{i_0-1}, u_{i_0}) < u'_{i_0} - \mu(u'_{i_0-1}, u'_{i_0})$. Since the function $f(\cdot)$ is strictly increasing, we obtain $f(u_{i_0} - \mu(u_{i_0-1}, u_{i_0})) < f(u'_{i_0} - \mu(u'_{i_0-1}, u'_{i_0}))$. Since $\alpha_{i_0,i_0-1} > 0$ we further have $\gamma^{(1)}_{i_0}(\mathbf{u}) < \gamma^{(1)}_{i_0}(\mathbf{u}')$. Applying inductively Lemma 2 (note that at each application, $L2$ holds) establishes condition $C(M-1)$ and

$$\gamma^{(1)}_{M-1}(\mathbf{u}) < \gamma^{(1)}_{M-1}(\mathbf{u}'). \tag{20}$$

On the other side, since $u_{M-1} < u'_{M-1}$, it follows from $T2$ that $\mu(u_{M-1}, W) - u_{M-1} \geq \mu(u'_{M-1}, W) - u'_{M-1}$. Hence $f(\mu(u_{M-1}, W) - u_{M-1}) \geq f(\mu(u'_{M-1}, W) - u'_{M-1})$. It follows that

$$\gamma^{(2)}_{M-1}(\mathbf{u}) \geq \gamma^{(2)}_{M-1}(\mathbf{u}'). \tag{21}$$

Relations (20), (21) together with $\gamma^{(1)}_{M-1}(\mathbf{u}') = \gamma^{(2)}_{M-1}(\mathbf{u}')$ and $\gamma^{(1)}_{M-1}(\mathbf{u}) = \gamma^{(2)}_{M-1}(\mathbf{u})$ lead to a contradiction.

Case 2. $i_0 = 1$. Applying inductively Lemma 2 concludes that at least one of the following two assertions holds:

$A1$) $C(M-1)$ is satisfied and $\gamma^{(1)}_{M-1}(\mathbf{u}) < \gamma^{(1)}_{M-1}(\mathbf{u}')$.

$A2$) $C(i)$ is satisfied for all $i, 1 \leq i \leq M-1$. Additionally, $\mu(u_i, u_{i+1}) - u_i = \mu(u'_i, u'_{i+1}) - u'_i$ for all $i, 1 \leq i \leq M-2$.

If $A1$ is true, then a contradiction arises as in the previous case. If the second assertion is true, then we can apply $T4$ and it follows that at least one of the following statements is valid:

$A2.1$) $\mu(u_{M-1}, u_M) - u_{M-1} > \mu(u'_{M-1}, u'_M) - u'_{M-1}$ (by $T4.1$, since $u_M = u'_M = W$);

$A2.2$) there is some $i_1, 0 \leq i_1 \leq M-2$, such that $u_{i_1+1} - \mu(u_{i_1}, u_{i_1+1}) < u'_{i_1+1} - \mu(u'_{i_1}, u'_{i_1+1})$ (by $T4.3$ and $T4.1$; note that $u_0 = u'_0 = V$).

When $A2.1$ holds, we have $\gamma^{(2)}_{M-1}(\mathbf{u}) > \gamma^{(2)}_{M-1}(\mathbf{u}')$ because $f(\cdot)$ is strictly increasing and $\beta_{M-1,M} > 0$. Using (10) we obtain that $\gamma^{(1)}_{M-1}(\mathbf{u}) > \gamma^{(1)}_{M-1}(\mathbf{u}')$. On the other hand, because $C(M-1)$ is satisfied it follows that $\gamma^{(1)}_{M-1}(\mathbf{u}) \leq \gamma^{(1)}_{M-1}(\mathbf{u}')$ (by the argument used to derive (17) in the proof of Lemma 2). Thus, we have reached a contradiction.

Now we treat the case when $A2.2$ holds. Note first that since $C(i_1+1)$ is satisfied, by (16) we have

$$f(u_{i_1+1} - \mu(u_j, u_{i_1+1})) \leq f(u'_{i_1+1} - \mu(u'_j, u'_{i_1+1})) \text{ for all } j, 0 \leq j < i_1 + 1, \tag{22}$$

as shown in the proof of Lemma 2. Moreover, the inequality (22) corresponding to $j = i_1$ is strict due to the condition in $A2.2$ and the strict monotonicity of $f(\cdot)$. Further, because all $\alpha_{i_1+1,j}$ are non-negative and $\alpha_{i_1+1,i_1} > 0$, it follows that $\gamma^{(1)}_{i_1+1}(\mathbf{u}) < \gamma^{(1)}_{i_1+1}(\mathbf{u}')$. The same argument as in Case 1 leads to a contradiction as well. □

We say that an IA is optimal if there is a globally optimal convex K-DSQ (globally optimal with respect to the set of convex K-DSQs), which has that IA. The following result is a direct consequence of Theorem 1.

Corollary 1. If Conditions A, and B hold, $\mathbf{Q}$ is a locally optimal convex K-DSQ, its IA is optimal, and the conditions in Theorem 1 are satisfied (e.g., if $p(x)$ is log-concave and the error function is the squared difference [24, Theorem 4]), then $\mathbf{Q}$ is a globally optimal convex K-DSQ.

Consequently, when the sufficient conditions of Theorem 1 are met, the design of optimal convex K-DSQ can be performed by first finding an optimal IA, then applying a generalized Lloyd-type algorithm to optimize the K-DSQ, given that assignment.

Finding the optimal IA is a difficult problem in general. Fortunately, for convex MDSQ's there are several important cases when this problem is easier to handle, such as MRSQ and symmetric MDSQ. In the next two sections we address the problem of IA for these two cases.

V. Optimal Index Assignment for Convex MRSQ

This section is devoted to the discussion of optimal IA for convex MRSQ. Recall that a convex MRSQ with K refinement stages is a convex K-DSQ whose active components are $Q_1, Q_{\{1,2\}}, \cdots, Q_{\{1,\cdots,i\}}, \cdots, Q_{\mathcal{K}}$, where $Q_{\{1,\cdots,i\}}$ denotes the component quantizer corresponding to the first i descriptions, hence its partition is the intersection of the partitions of side quantizers $Q_1, Q_2, \cdots, Q_i$. In our definition of convex K-DSQ we have imposed only to active components to have convex cells. Therefore, in a convex MRSQ only quantizers $Q_1, Q_{\{1,2\}}, \cdots, Q_{\{1,\cdots,i\}}, \cdots, Q_{\mathcal{K}}$ are required to satisfy this constraint, while the side quantizers $Q_2, Q_3, \cdots, Q_K$ may have nonconvex cells.

Interestingly, as the following theorem shows, in the case of convex MRSQ, only the number of cells in the central partition is relevant at optimality, and not the IA. The intuitive reason for the theorem to hold is that if the central partition has maximal number of cells, then it determines all active side quantizers.

Theorem 2. Assuming that Conditions A and B hold, a globally optimal convex MRSQ of K refinement stages must have $M = M_1 M_2 \cdots M_K$ cells in the central partition, where M_i denotes the number of cells in side quantizer Q_i, for all $1 \leq i \leq K$. As long as the latter condition is satisfied any index assignment is optimal.

Proof. In order to prove this theorem it is useful to note that the active components of the convex MRSQ $Q_1, Q_{\{1,2\}}, \cdots, Q_{\{1,2,\cdots,i-1,i\}}, \cdots, Q_{\mathcal{K}}$ forms a sequence of embedded convex quantizers, i.e., any cell of $Q_{\{1,2,\cdots,i-1\}}$ is the union of some cells of $Q_{\{1,2,\cdots,i-1,i\}}$, more specifically, of at most M_i such cells. Another way of saying this is that the partition of $Q_{\{1,2,\cdots,i-1,i\}}$ is obtained by splitting each cell of $Q_{\{1,2,\cdots,i-1\}}$ into at most M_i nonempty subintervals. To see this, let C be a cell of $Q_{\{1,\cdots,i-1\}}$ and let $C_1^{(i)}, C_2^{(i)}, \cdots, C_{M_i}^{(i)}$, be the cells of the side quantizer Q_i. Then the sets $C \cap C_1^{(i)}, C \cap C_2^{(i)}, \cdots, C \cap C_{M_i}^{(i)}$ (actually those which are nonempty) are cells of $Q_{\{1,2,\cdots,i-1,i\}}$ and

$$C = (C \cap C_1^{(i)}) \cup (C \cap C_2^{(i)}) \cup \cdots \cup (C \cap C_{M_i}^{(i)}).$$

Further we argue that at optimality all the sets $C \cap C_1^{(i)}, C \cap C_2^{(i)}, \cdots, C \cap C_{M_i}^{(i)}$ have to be non-empty. In order to prove this point it is enough to show that if one of these sets is empty then a convex MRSQ of strictly smaller expected distortion can be constructed. For this, assume that $C \cap C_1^{(i)} = \emptyset$ and $C \cap C_2^{(i)} = (a, b] \neq \emptyset$. Pick some point t inside the open interval (a, b) and define two new sets $A_1^{(i)} = C_1^{(i)} \cup (a, t]$ and $A_2^{(i)} = C_2^{(i)} - (a, t]$. The new MRSQ is constructed by replacing the cells $C_1^{(i)}, C_2^{(i)}$ of side quantizer Q_i by the sets $A_1^{(i)}, A_2^{(i)}$, respectively, and optimizing the MRSQ decoder. We assume that the decoder of the old MRSQ was optimized as well. Let us analyze now the effect of this change on the active components of the MRSQ. Clearly, quantizers $Q_1, Q_{\{1,2\}}, \cdots, Q_{\{1,2,\cdots,i-1\}}$ are not affected. The only effect on the partition of $Q_{\{1,2,\cdots,i-1,i\}}$ is that the old cell $(a, b]$ is replaced by two new non-empty cells $(a, t] = C \cap A_1^{(i)}$ and $(t, b] = C \cap A_2^{(i)}$. For $j, i + 1 \leq j \leq K$, only the portion of the partition of $Q_{\{1,2,\cdots,j\}}$ covering the interval $(a, b]$ is affected. Note that in the old MRSQ this portion of the partition consists of all non-empty intersections $(a, b] \cap C'$, where C' ranges over all non-empty intersections of cells of side quantizers $Q_{i+1}, \cdots, Q_j$. In the new MRSQ this portion of the partition is composed of all non-empty intersections $(a,t] \cap C'$ and $(t,b] \cap C'$ with all possible C' as above. Let $(a,b] \cap C' = (c,d]$. Then exactly one of the following three cases is possible: 1) $(a,t] \cap C' = \emptyset$ and $(t,b] \cap C' = (c,d]$ when $t \leq c$; 2) $(a,t] \cap C' = (c,t]$ and $(t,b] \cap C' = (t,d]$ when $c < t < d$; 3) $(a,t] \cap C' = (c,d]$ and $(t,b] \cap C' = \emptyset$ when $t \geq d$. Consequently, the interval $(c,d]$ from the old partition either remains unchanged in the new partition or is split into two non-empty intervals $(c,t]$ and $(t,d]$. Because there is only one cell $(c,d]$ which can contain t in its interior, it follows that at most one of the old cells is split. The above analysis reveals that the new MRSQ is still convex. Moreover, each of its active component quantizers is either identical to the old one, or is obtained from the old one by splitting one cell into two non-empty intervals, and at least one of the active components is in the second category, i.e., $Q_{\{1,2,\cdots,i\}}$. Since the pdf $p(x)$ is strictly positive, by splitting a cell of a convex quantizer into two non-empty intervals a convex quantizer of strictly smaller distortion is obtained. It follows that the new MRSQ has a strictly smaller distortion than the old one.

The above argument implies that in the optimal convex MRSQ, the partition of $Q_{\{1,2,\cdots,i-1,i\}}$ is obtained by splitting each cell of $Q_{\{1,2,\cdots,i-1\}}$ into exactly M_i nonempty subintervals. This implies that it must have exactly $M = M_1 M_2 \cdots M_K$ cells in the central partition. Now let us consider a convex MRSQ which has $M = M_1 M_2 \cdots M_K$ cells in the central partition and let us examine the relevance of the IA. Note that, since the number of cells in the central partition is maximal, then the central partition determines the partitions of the active components, irrespective of the IA. Precisely, for each $i, 1 \leq i \leq K-1$, each cell of quantizer $Q_{\{1,2,\cdots,i-1,i\}}$ must be the union of $M'_i = \Pi_{j=i+1}^{K} M_j$ consecutive intervals of the central partition. Therefore, if $u_0, u_1, u_2, \cdots u_M$ are the thresholds of the central partition, then the thresholds of the quantizer $Q_{\{1,2,\cdots,i-1,i\}}$ are necessarily $u_0, u_{M'_i}, u_{2M'_i}, \cdots, u_{(M_1 M_2 \cdots M_i - 1)M'_i}, u_M$. There is a multitude of IA's which yield the above partitions, but they affect only the distortions of non-active components and therefore they do not have an impact on the expected distortion of the MRSQ. We conclude that when the number of cells in the central partition is maximal,i.e., equals $M_1 M_2 \cdots M_K$, any index assignment is optimal. This observation concludes the proof. $\square$

The next corollary is an immediate consequence of Theorem 2 and Corollary 1 from the previous section.

Corollary 2. If Conditions A and B are satisfied, then any locally optimal convex MRSQ with $M_1 M_2 \cdots M_K$

number of cells in the central partition is globally optimal, too, if the conditions in Theorem 1 are satisfied.

The generalized Lloyd-type algorithm for MRSQ design proposed by [4] constructs a locally optimal (fixed-rate convex) MRSQ with the maximal number of cells in the central quantizer. According to the above corollary, if the conditions in Theorem 1 are satisfied, then the MRSQ obtained is globally optimal.

VI. Optimal Index Assignment for Symmetric Convex K-DSQ.

In this section we settle the problem of optimal IA for symmetric convex K-DSQ.

Recall that in a symmetric convex K-DSQ all side quantizers have the same rate, i.e., the same number of cells, hence $M_1 = M_2 = \cdots = M_K$. Moreover, the weight $\omega_{\mathcal{I}}$ is only a function of the cardinality of the set $\mathcal{I}$, in other words we have $\omega_{\mathcal{I}} = \omega_{\mathcal{I}'}$ if $|\mathcal{I}| = |\mathcal{I}'|$. We also require that $\omega_1 > 0$. Hence all side quantizers are active components, and according to our definition of convex K-DSQ they have convex cells. This implies that all component quantizers, active or not, are convex too.

Each side partition is specified by $M_1 - 1$ thresholds. Let $v_i^0 = V$, $v_i^{M_1} = W$ and let v_i^j, $0 \leq j \leq M_1$ satisfying

$$V = v_i^0 < v_i^1 < \cdots v_i^j < v_i^{j+1} < \cdots < v_i^{M_1} = W,$$

be the thresholds of side quantizer Q_i, $1 \leq i \leq K$.

The set of thresholds of the central quantizer is the union of the sets of threshold of all side quantizers. Therefore, the maximal number of cells in the central partition is $K(M_1 - 1) + 1$, and it is achieved if and only if the thresholds of different side partitions are different. Clearly, specifying an IA is equivalent to specifying the order of thresholds v_i^j in the central partition, and specifying the equalities between these thresholds, if any.

In this section we will consider only convex K-DSQ's with optimized decoders. Hence the representation point of any cell $(a, b]$ will be its generalized centroid $\mu(a, b)$. We define the distortion of the cell, denoted by $D(a, b)$, according to

$$D(a, b) = \int_a^b f(|x - \mu(a, b)|)p(x)dx.$$

Thus, the distortion of a quantizer becomes the sum of distortions of its cells. As previously, the error function $f(\cdot)$ is assumed to be convex, continuous and strictly increasing. We will also prove some results for the case when $f(\cdot)$ satisfies additional requirements, and then these requirements will be specified.

Some progress toward finding the optimal IA for symmetric convex 2-DSQ was achieved in [7]. Precisely, it was proved that there exists an optimal symmetric convex 2-DSQ which satisfies the following inequalities

$$v_1^1 \leq v_2^1 \leq v_1^2 \leq \cdots v_1^j \leq v_2^j \leq v_1^{j+1} \leq \cdots \leq v_2^{M_1-1}. \tag{23}$$

While this result sheds considerable light on the structure of the optimal IA in the case of two symmetric descriptions, it does not solve the problem completely. In order to uniquely identify an IA each inequality "$\leq$" above must be replaced by "$<$" or by "$=$". Therefore, the series of relations (23) characterize a class of at least 2^{M_1-1} distinct IA's. This count was obtained by considering all possibilities when every second inequality in (23) is strict and any other inequality is replaced by "$<$" or by "$=$". Consequently, in order to find the optimal IA, a search must be conducted among at least 2^{M_1-1} possibilities. Nevertheless, the structure highlighted by relations 23) was exploited in [7] to accelerate the combinatorial design algorithm. The case of $K > 2$ was not treated in [7] and the proof given there for $K = 2$ does not extend in a straightforward manner to more than two descriptions.

In this section we settle the problem of optimal IA for symmetric convex K-DSQ, for general K. Specifically, we prove that there is an optimal IA for which a series of relations similar to (23) hold, but with strict inequalities. Consequently, these relations uniquely specify an IA. Furthermore, we prove that when the error function $f(\cdot)$ is additionally continuously differentiable, any optimal index assignment must satisfy these relations, up to a permutation of the side quantizers. The first step toward our main result is the next theorem which clarifies that an optimal convex K-DSQ whose all side quantizers are active components (and hence convex), must have the maximal possible number of cells in the central partition, as allowed by the side quantizers rates. Note that the result established by this proposition is not restricted to symmetric K-DSQ's, in other words, the side quantizers may have different rates and different weights.

Theorem 3. Assume that Conditions A and B hold. Consider the class of convex K-DSQ's whose all side quantizers are active components, and each side quantizer Q_i has M_i cells, $1 \leq i \leq K$. Then there is an optimal K-DSQ within this class and it necessarily contains $M_1 + M_2 + \cdots + M_K - K + 1$ (non-empty) cells in the central partition.

Proof. Clearly, to find an optimal K-DSQ we only need to look at decoder optimized K-DSQ's. Consider first the decoder optimized K-DSQ's within the class specified in the hypothesis of the theorem, which have M cells in the central partition, for some M such that $\max_{1\le i\le K} M_i \le M \le M_1+M_2+\cdots+M_K-K+1$, and some IA denoted by h_M. Let $\mathcal{O}_M$ be the set of $(M-1)$-dimensional real vectors $\mathbf{u}=(u_1,u_2,\cdots,u_{M-1})$ satisfying $V<u_1<u_2<\cdots<u_{M-1}<W$. Define the function $F_{h_M}(\mathbf{u})$ on the set $\mathcal{O}_M$ as follows

$$F_{h_M}(\mathbf{u}) = \sum_{\mathcal{I}\subseteq\mathcal{K},\omega_{\mathcal{I}}\neq 0} \omega_{\mathcal{I}} \sum_{k=1}^{M_{\mathcal{I}}} \int_{u_{j_{\mathcal{I},k-1}}}^{u_{j_{\mathcal{I},k}}} f(|x-\mu(u_{j_{\mathcal{I},k-1}},u_{j_{\mathcal{I},k}})|)p(x)dx, \tag{24}$$

with the notations introduced in Section III. Then $F_{h_M}(\mathbf{u})$ represents the expected distortion of the convex K-DSQ with vector of thresholds $\mathbf{u}$ and with h_M as IA. Next, denote by $\overline{\mathbf{R}}$ the extended real line and let $\overline{\mathcal{O}}_M$ be the set of all $(u_1,u_2,\cdots,u_{M-1})\in\overline{\mathbf{R}}^{M-1}$ satisfying $V\le u_1\le u_2\le\cdots\le u_{M-1}\le W$. Consider on the set $\overline{\mathcal{O}}_M$ the topology inherited from the product topology on $\overline{\mathbf{R}}^{M-1}$. Now extend the function $F_{h_M}(\mathbf{u})$ on the set $\overline{\mathcal{O}}_M$ by replacing $\int_{u_{j_{\mathcal{I},k-1}}}^{u_{j_{\mathcal{I},k}}} f(|x-\mu(u_{j_{\mathcal{I},k-1}},u_{j_{\mathcal{I},k}})|)p(x)dx$ in (24) by 0, whenever $u_{j_{\mathcal{I},k-1}}=u_{j_{\mathcal{I},k}}$. By similar arguments to those used by Kieffer in [18, Lemma 3] and [19, Lemma 2], the function $F_{h_M}(\cdot)$ is continuous on $\overline{\mathbf{R}}^{M-1}$. Since $\overline{\mathbf{R}}^{M-1}$ is a compact set it follows that the function $F_{h_M}(\cdot)$ achieves its minimum on $\overline{\mathbf{R}}^{M-1}$. Let $\hat{\mathbf{u}}_{h_M}\in\overline{\mathcal{O}}_M$ denote such a minimum point, i.e. $F_{h_M}(\hat{\mathbf{u}}_{h_M})=\min_{\mathbf{u}\in\overline{\mathcal{O}}_M} F_{h_M}(\mathbf{u})$. Next, let M_0 and h_{M_0} be such that $F_{h_{M_0}}(\hat{\mathbf{u}}_{h_{M_0}})=\min_{M,h_M} F_{h_M}(\hat{\mathbf{u}}_{h_M})$, where the minimum is taken over all possible numbers M of cells in the central partition and IA's h_M corresponding to the class of convex K-DSQ's whose all side quantizers are active components, and each side quantizer Q_i has M_i cells. Clearly, the minimum exists since the set of all possible such IA's is finite. Moreover, $F_{h_{M_0}}(\hat{\mathbf{u}}_{h_{M_0}})$ is smaller or equal than the distortion of any convex K-DSQ in the class specified above. Therefore, in order to complete the proof it is enough to show that $M_0=M_1+M_2+\cdots+M_K-K+1$ and that $\hat{\mathbf{u}}_{h_{M_0}}\in\mathcal{O}_{M_0}$.

Note that $\hat{\mathbf{u}}_{h_{M_0}}$ can be interpreted as the vector of thresholds of a convex K-DSQ $\mathbf{Q}$ with all side quantizers being active components, where the number of cells in Q_i is at most M_i and such that $\bar{D}(\mathbf{Q})=F_{h_{M_0}}(\hat{\mathbf{u}}_{h_{M_0}})$. If for some i's, $1\le i\le K$, the number of cells of Q_i were less than M_i, then by refining the encoder partition of each such Q_i to have exactly M_i non-empty cells we would obtain a new K-DSQ (with optimized decoder) of expected distortion strictly lower (because all $\omega_i>0$ and the pdf $p(x)$ is strictly positive), fact which leads to a contradiction. It follows that each Q_i must have exactly M_i cells. In order to complete the proof is then enough to show that no two side quantizers have thresholds in common.

We will present a proof by contradiction for the above claim. Note that all since all side quantizers are active components, it follows that they are all convex, hence all component quantizers, active or not, are convex too. Assume without restricting the generality that the first k side quantizers, for some $k \geq 2$, have a threshold in common. Then there are $j_1, \cdots, j_k$ with $1 \leq j_i \leq M_i - 1$, and v such that $v_1^{j_1} = \cdots = v_k^{j_k} = v$ and $v_i^j \neq v$ for any $i > k$ and any j. We will construct a new convex K-DSQ $\mathbf{Q}'$ starting from $\mathbf{Q}$, such that $\bar{D}(\mathbf{Q}') < \bar{D}(\mathbf{Q})$. The construction is based on a properly chosen perturbation of the threshold v in Q_k. For this we need to introduce first some notations. For each $\mathcal{I} \subseteq \mathcal{K}$, such that $\{1, 2, \cdots, k\} \cap \mathcal{I} \neq \emptyset$, let $v_{\mathcal{I}}^l$, respectively $v_{\mathcal{I}}^r$, denote the threshold preceding, respectively following, v in the encoder partition of $Q_{\mathcal{I}}$ (note that superscript l stands for left and superscript r stands for right). Further, let $y_{\mathcal{I}}^l$ denote the reconstruction value (hence generalized centroid) of cell $(v_{\mathcal{I}}^l, v]$, and $y_{\mathcal{I}}^r$ denote the reconstruction value of cell $(v, v_{\mathcal{I}}^r]$. Recall that the generalized centroid is contained in the interior of the cell. Also denote by $\mathcal{S}_1$ the set of all non-empty subsets $\mathcal{I}$ of $\mathcal{K}$ such that $k \in \mathcal{I}$ and $\{1, 2, \cdots, k-1\} \cap \mathcal{I} = \emptyset$. Also, let $\mathcal{S}_2$ denote the set of all non-empty subsets $\mathcal{I}$ of $\mathcal{K}$ such that $k \in \mathcal{I}$ and $\{1, 2, \cdots, k-1\} \cap \mathcal{I} \neq \emptyset$. Next we need to distinguish between two cases. The first case is when the following inequality holds

$$\sum_{\mathcal{I} \in \mathcal{S}_1} \omega_{\mathcal{I}} \big(f(|v - y_{\mathcal{I}}^r|) - f(|v - y_{\mathcal{I}}^l|) \big) \geq 0. \tag{25}$$

Let $y_0 = \min_{\mathcal{J} \in \mathcal{S}_2} y_{\mathcal{J}}^r$. Clearly, we have $v < y_0$. Further, let $y = \frac{v + y_0}{2}$. Hence $v < y < y_0$. These relations together with the definition of y_0 and the fact that $f(\cdot)$ is strictly increasing, imply that $f(|v - y_{\mathcal{J}}^r|) > f(|v - y|)$ for any $\mathcal{J} \in \mathcal{S}_2$. Using (25) and the fact that there is some $\mathcal{J} \in \mathcal{S}_2$ such that $\omega_{\mathcal{J}} > 0$ (precisely, $\mathcal{J} = \{k\}$), we conclude that the following inequality holds

$$\sum_{\mathcal{I} \in \mathcal{S}_1} \omega_{\mathcal{I}} \big(f(|v - y_{\mathcal{I}}^r|) - f(|v - y_{\mathcal{I}}^l|) \big) + \sum_{\mathcal{J} \in \mathcal{S}_2} \omega_{\mathcal{J}} \big(f(|v - y_{\mathcal{J}}^r|) - f(|v - y|) \big) > 0. \tag{26}$$

Consider now the function $g(x)$ defined for $x \in [v, y]$ as follows

$$g(x) = \sum_{\mathcal{I} \in \mathcal{S}_1} \omega_{\mathcal{I}} \big(f(|x - y_{\mathcal{I}}^r|) - f(|x - y_{\mathcal{I}}^l|) \big) + \sum_{\mathcal{J} \in \mathcal{S}_2} \omega_{\mathcal{J}} \big(f(|x - y_{\mathcal{J}}^r|) - f(|x - y|) \big).$$

Equation (26) implies that $g(v) > 0$, and since g is continuous, it follows that there is some $\delta > 0$ with $v + \delta \leq y$, such that $g(x) > 0$ for all $x \in [v, v + \delta]$. Finally, let u denote the threshold following v in the central partition of $\mathbf{Q}$, and define $v' = \min(v + \delta, \frac{v+u}{2})$, hence $v' < u$. Now we are ready to construct the new convex K-DSQ $\mathbf{Q}'$ starting from $\mathbf{Q}$. For this we replace the threshold $v_k^{j_k}$ in side quantizer Q_k by v' and keep all other thresholds fixed. Note that this change does not affect the quantizers $Q_{\mathcal{I}}$ for subsets $\mathcal{I}$ which do not contain k. For $\mathcal{I} \in \mathcal{S}_1$ only two cells are affected. Precisely, cells $(v_{\mathcal{I}}^l, v]$, $(v, v_{\mathcal{I}}^r]$ are changed into $(v_{\mathcal{I}}^l, v']$, $(v', v_{\mathcal{I}}^r]$, respectively. Moreover, the effect on quantizers $Q_{\mathcal{J}}$ with $\mathcal{J} \in \mathcal{S}_2$ is that one cell is split into two. Specifically, cell $(v, v_{\mathcal{J}}^r]$ is split into $(v, v']$ and $(v', v_{\mathcal{J}}^r]$.

Now in order to completely characterize the new K-DSQ $\mathbf{Q}'$ we have to specify its reconstruction values as well. For all cells which have not been changed, we keep the same reconstruction values as in $\mathbf{Q}$. Further, for quantizers $Q_{\mathcal{I}}$ with $\mathcal{I} \in \mathcal{S}_1$, we let $y_{\mathcal{I}}^l$, respectively $y_{\mathcal{I}}^r$, be the reconstruction values of cells $(v_{\mathcal{I}}^l, v']$, $(v', v_{\mathcal{I}}^r]$, respectively. Finally, for quantizers $Q_{\mathcal{J}}$ with $\mathcal{J} \in \mathcal{S}_2$ we let y, respectively $y_{\mathcal{J}}^r$, be the reproduction value of cell $(v, v']$, respectively $(v', v_{\mathcal{J}}^r]$. Now it is clear that the change from $\mathbf{Q}$ to $\mathbf{Q}'$ incurs a change in the mapping of source samples to reproduction values only for the samples x in $(v, v']$ and only for subsets $\mathcal{I}$ of descriptions with $\mathcal{I} \in \mathcal{S}_1 \cup \mathcal{S}_2$. Thus we obtain the following equality

$$\bar{D}(\mathbf{Q}) - \bar{D}(\mathbf{Q}') = \int_v^{v'} \Big(\sum_{\mathcal{I} \in \mathcal{S}_1} \omega_{\mathcal{I}} \big(f(|x - y_{\mathcal{I}}^r|) - f(|x - y_{\mathcal{I}}^l|) \big) + \sum_{\mathcal{J} \in \mathcal{S}_2} \omega_{\mathcal{J}} \big(f(|x - y_{\mathcal{J}}^r|) - f(|x - y|) \big) \Big) p(x) dx.$$

It follows that

$$\bar{D}(\mathbf{Q}) - \bar{D}(\mathbf{Q}') = \int_v^{v'} g(x) p(x) dx.$$

The definition of v' implies that $g(x) > 0$ for all $x \in (v, v')$. Moreover, since $p(x) > 0$ for $x \in (v, v')$, too, we obtain that $g(x)p(x) > 0$ for all $x \in (v, v')$, and further that $\int_v^{v'} g(x)p(x)dx > 0$. This leads to the conclusion that $\bar{D}(\mathbf{Q}') < \bar{D}(\mathbf{Q})$, which contradicts the optimality of $\mathbf{Q}$. Thus, the proof of this case is completed. The case when relation (25) does not hold can be treated symmetrically by appropriately choosing a value $v' < v$ and constructing $\mathbf{Q}'$ by replacing $v_k^{j_k}$ by v'. $\square$

Remark 1. As an immediate corollary to the above proposition, it follows that all inequalities in (23) are strict, and hence they define an optimal IA for symmetric 2-DSQ.

The next results establishes an optimal IA for symmetric convex K-DSQ, for general K.

Theorem 4. Assume that Conditions A and B hold. Then there is an optimal symmetric convex K-DSQ such that

$$v_1^1 < v_2^1 < \cdots < v_i^j < v_{i+1}^j < \cdots < v_K^j < v_1^{j+1} < \cdots < v_K^{M_1-1}, \tag{27}$$

in other words, where $v_i^j < v_{i+1}^j$ holds for all $1 \le i \le K-1$ and $1 \le j \le M_1 - 1$, and $v_K^j < v_1^{j+1}$ holds for all $1 \le j \le M_1 - 2$. Moreover, if $f(\cdot)$ is additionally continuously differentiable, then any optimal symmetric convex K-DSQ must satisfy the relations (27) possibly after a permutation of subscripts of the side quantizers.

Proof. Note that by Theorem 3, an optimal symmetric convex K-DSQ with M_1 cells in each side quantizer must exist. Further, the idea of the proof is to show that by exchanging thresholds between side quantizers such that (27) to be satisfied, the expected distortion of the K-DSQ does not increase. It will be understood that after performing such an operation the decoder will be optimized, and we will not explicitly state this.

First we permute the entire partitions among side quantizers (or equivalently, apply a permutation of the side quantizers subscripts) such that $v_1^1 < v_2^1 < \cdots < v_K^1$, in other words such that the first $K-1$ inequalities in (27) to be satisfied. This permutation results in a convex K-DSQ with the same expected distortion. Further, let us order of thresholds $v_i^j, 1 \le i \le K, 1 \le j \le M_1 - 1$, in increasing order. Note that according to Theorem 3, the elements of this sequence are pairwise distinct. Denote by t_l the l-th element in the ordered sequence, for $1 \le l \le K(M_1-1)$. Further, let $\prec$ denote the lexicographical order (l.o., for short) of pairs of integers (j,i). Precisely, $(j,i) \prec (j',i')$ if and only if either 1) $j < j'$ or 2) $j = j'$ and $i < i'$. If $(j,i) \prec (j',i')$ we say that the pair (j,i) is smaller than the pair (j',i') in l.o.. Note that the ordering of thresholds v_i^j in (27) corresponds to the lexicographical order of the pairs (j,i). Now let us order the pairs (j,i), $1 \le i \le K, 1 \le j \le M_1 - 1$, in l.o., and denote by $o(j,i)$ the position of pair (j,i) in this sequence. Then $o(j,i) = (j-1)K + i$ for all $1 \le i \le K, 1 \le j \le M_1 - 1$. Clearly, the inequalities (27) are satisfied if

$$v_i^j = t_{o(j,i)} \tag{28}$$

for all $1 \le i \le K, 1 \le j \le M_1 - 1$.

To make the explanation more intuitive we say that the threshold v_i^j is correctly placed if the above equality is satisfied, and we say that it is misplaced, otherwise. Notice that, due to the permutation of quantizer subscripts applied earlier, the first K thresholds in sequence are correctly placed. We will exchange thresholds between side quantizers in a series of steps such that after each step the number of correct placements in the sequence of thresholds up to the first misplacement strictly increases, while the expected distortion does not decrease. Assume that we are at the beginning of some step $s, s \geq 1$, and that the first misplacement occurs in the ℓ-th position, for some $\ell > K$. In other words, equality (28) holds for all pairs (j, i) such that $o(j, i) < \ell$ and does not hold for the pair (j_1, i_1) satisfying $o(j_1, i_1) = \ell$. Let (j_2, i_2) be the pair for which $t_\ell = v_{i_2}^{j_2}$. Then clearly, $j_1 > 1$ and the following relations are valid

$$(j_1, i_1) \prec (j_2, i_2) \tag{29}$$

and

$$v_{i_2}^{j_2} < v_{i_1}^{j_1}.$$

Since $v_{i_2}^{j_2-1} < v_{i_2}^{j_2}$, it follows that $v_{i_2}^{j_2-1} = t_{\ell'}$ for some $\ell' < \ell$, hence $v_{i_2}^{j_2-1}$ is correctly placed. This implies that $o(j_2-1, i_2) = \ell' \leq \ell-1$. Moreover, $v_{i_1}^{j_1-1}$ is also correctly placed because $o(j_1-1, i_1) < o(j_1, i_1) = \ell$. Inequality (29) implies that $o(j_1 - 1, i_1) < o(j_2 - 1, i_2)$ (note that necessarily $j_2 > 1$ since $j_1 > 1$). Furthermore, since both $v_{i_1}^{j_1-1}$ and $v_{i_2}^{j_2-1}$ are correctly placed, we obtain that

$$v_{i_1}^{j_1-1} < v_{i_2}^{j_2-1}$$

Summarizing, we have established the following sequence of inequalities, which is crucial to our development

$$v_{i_1}^{j_1-1} < v_{i_2}^{j_2-1} < v_{i_2}^{j_2} < v_{i_1}^{j_1}. \tag{30}$$

In order to describe the interchanges that we will make, let k be the smallest nonnegative integer such that $k \leq M_1 - j_2$ and $v_{i_1}^{j_1+k} \leq v_{i_2}^{j_2+k}$. Such an integer always exists and is strictly positive because the previous inequality is satisfied for $k = M_1 - j_2$ (since $v_{i_2}^{M_1} = W$) and is not satisfied for $k = 0$ by (30). Then the following sequence of inequalities holds:

$$v_{i_2}^{j_2+k-1} < v_{i_1}^{j_1+k-1} < v_{i_1}^{j_1+k} \leq v_{i_2}^{j_2+k}. \tag{31}$$

Notice that we have $v_{i_1}^{j_1+k} = v_{i_2}^{j_2+k}$ only if $j_1+k = j_2+k = M_1$ (by Theorem 3), otherwise the inequality is strict. Now interchange the thresholds $v_{i_2}^{j_2}, \cdots, v_{i_2}^{j_2+k-1}$, with $v_{i_1}^{j_1}, \cdots, v_{i_1}^{j_1+k-1}$, respectively, between the side partitions Q_{i_2} and Q_{i_1}. This interchange is illustrated in Figure 3. Note that the number of cells in the two side quantizers is not affected by this operation. Moreover, this interchange does not affect the first $\ell-1=o(j_1,i_1)-1$ thresholds in the sequence, and causes the threshold on position ℓ to become correctly placed, too. Thus, the number of correctly placed thresholds up to the first misplacement strictly increases. It remains to show that the expected distortion of the K-DSQ does not increase. For this we will consider pairs of component quantizers and analyze how their contribution to the expected distortion is affected.

Let us first consider the side quantizers Q_{i_2} and Q_{i_1}. The partition of Q_{i_2} is modified only between thresholds $v_{i_2}^{j_2-1}$ and $v_{i_2}^{j_2+k}$. The new cells are:

$$(v_{i_2}^{j_2-1},v_{i_1}^{j_1}],(v_{i_1}^{j_1},v_{i_1}^{j_1+1}],\cdots,(v_{i_1}^{j_1+k-2},v_{i_1}^{j_1+k-1}],(v_{i_1}^{j_1+k-1},v_{i_2}^{j_2+k}].$$

The partition of Q_{i_1} is modified only between $v_{i_1}^{j_1-1}$ and $v_{i_1}^{j_1+k}$, the new cells being

$$(v_{i_1}^{j_1-1},v_{i_2}^{j_2}],(v_{i_2}^{j_2},v_{i_2}^{j_2+1}],\cdots,(v_{i_2}^{j_2+k-2},v_{i_2}^{j_2+k-1}],(v_{i_2}^{j_2+k-1},v_{i_1}^{j_1+k}].$$

Note that the cells $(v_{i_2}^{j_2},v_{i_2}^{j_2+1}],\cdots,(v_{i_2}^{j_2+k-2},v_{i_2}^{j_2+k-1}]$, and $(v_{i_1}^{j_1},v_{i_1}^{j_1+1}],\cdots,(v_{i_1}^{j_1+k-2},v_{i_1}^{j_1+k-1}]$ have simply been exchanged, respectively, between side quantizers Q_{i_2} and Q_{i_1}. Since the distortions of Q_{i_2} and Q_{i_1} are weighted equally in the expected distortion of the K-DSQ, and since the distortion of each quantizer is the sum of distortions of its cells, it follows that the exchange of cells does not affect the overall contribution of the two side quantizers to the expected distortion. Thus, any change in the expected distortion is due only to the modification of the old cells $(v_{i_2}^{j_2-1},v_{i_2}^{j_2}]$, $(v_{i_2}^{j_2+k-1},v_{i_2}^{j_2+k}]$ of Q_{i_2} into $(v_{i_2}^{j_2-1},v_{i_1}^{j_1}]$, $(v_{i_1}^{j_1+k-1},v_{i_2}^{j_2+k}]$ respectively, and of the old cells $(v_{i_1}^{j_1-1},v_{i_1}^{j_1}]$, $(v_{i_1}^{j_1+k-1},v_{i_1}^{j_1+k}]$ of Q_{i_1} into $(v_{i_1}^{j_1-1},v_{i_2}^{j_2}]$, $(v_{i_2}^{j_2+k-1},v_{i_1}^{j_1+k}]$ respectively.

Let Δ denote the difference in the expected distortion due to the changes in Q_{i_2} and Q_{i_1}. Then

$$\Delta=\omega_1[D(v_{i_2}^{j_2-1},v_{i_1}^{j_1})+D(v_{i_1}^{j_1-1},v_{i_2}^{j_2})-D(v_{i_2}^{j_2-1},v_{i_2}^{j_2})-D(v_{i_1}^{j_1-1},v_{i_1}^{j_1})+$$
$$+D(v_{i_1}^{j_1+k-1},v_{i_2}^{j_2+k})+D(v_{i_2}^{j_2+k-1},v_{i_1}^{j_1+k})-D(v_{i_2}^{j_2+k-1},v_{i_2}^{j_2+k})-D(v_{i_1}^{j_1+k-1},v_{i_1}^{j_1+k})]. \tag{32}$$

At this point we apply Lemma 3, which is stated and proved in Appendix. Thus, from (30) we obtain:

$$D(v_{i_2}^{j_2-1},v_{i_1}^{j_1})+D(v_{i_1}^{j_1-1},v_{i_2}^{j_2})-D(v_{i_2}^{j_2-1},v_{i_2}^{j_2})-D(v_{i_1}^{j_1-1},v_{i_1}^{j_1})\leq 0. \tag{33}$$

Further, from (31), by applying Lemma 3 we obtain:

$$D(v_{i_1}^{j_1+k-1},v_{i_2}^{j_2+k})+D(v_{i_2}^{j_2+k-1},v_{i_1}^{j_1+k})-D(v_{i_2}^{j_2+k-1},v_{i_2}^{j_2+k})-D(v_{i_1}^{j_1+k-1},v_{i_1}^{j_1+k})\leq 0. \tag{34}$$

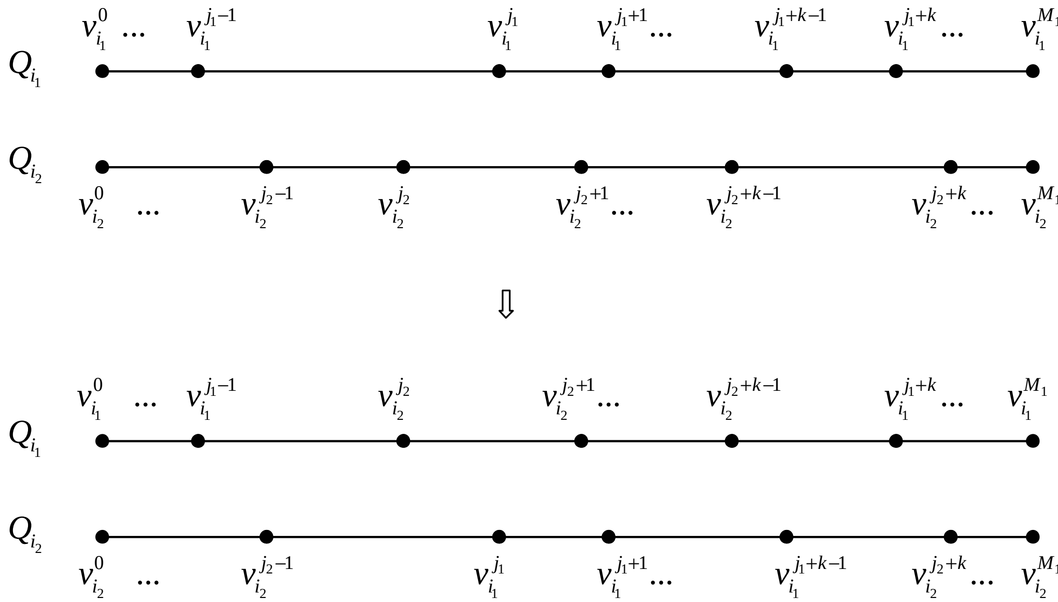

Fig. 3. Interchange of thresholds between side partitions Q_{i_2} and Q_{i_1}.

Relations (33) and (34) together with $\omega_1 > 0$, imply that $\Delta \leq 0$. Moreover, according to Lemma 3, when the error function $f(\cdot)$ is additionally continuously differentiable, the inequality (34) is strict, which implies that $\Delta < 0$.

Let us analyze now the modification incurred by the threshold interchange, on the other component quantizers. Clearly, any $Q_{\mathcal{I}}$ such that $i_1 \notin \mathcal{I}$, and $i_2 \notin \mathcal{I}$, is not affected. Also, the partition of $Q_{\{i_1,i_2\}}$ remains unchanged. Thus, the partition of any component quantizer $Q_{\mathcal{I}}$ such that $i_1, i_2 \in \mathcal{I}$ does not change either. Consider now an arbitrary $\mathcal{I} \subseteq \mathcal{K}$, $\mathcal{I} \neq \emptyset$, which does not contain either i_1 or i_2. We will analyze the changes incurred in $Q_{\{i_1\}\cup\mathcal{I}}$ and $Q_{\{i_2\}\cup\mathcal{I}}$. Consider the set $\mathcal{V}$ of thresholds t of quantizer $Q_{\mathcal{I}}$ such that $t \geq v_{i_2}^{j_2-1}$ and $t \leq v_{i_1}^{j_1+k}$. Assume first that $\mathcal{V}$ is non-empty and let $v = \min \mathcal{V}$ and $v' = \max \mathcal{V}$.

Case 1. $v < v_{i_2}^{j_2}$ and $v_{i_1}^{j_1+k-1} < v'$. This case is illustrated in Figure 4. The only effect in this case is that all cells between v and v' are exchanged between quantizers $Q_{\{i_2\}\cup\mathcal{I}}$ and $Q_{\{i_1\}\cup\mathcal{I}}$. Since the distortions of $Q_{\{i_2\}\cup\mathcal{I}}$ and $Q_{\{i_1\}\cup\mathcal{I}}$ are equally weighted in the total expected distortion, this exchange does not affect the expected distortion of the K-DSQ.

Case 2. $v_{i_2}^{j_2} < v$ and $v_{i_1}^{j_1+k-1} < v'$. Figure 5 illustrates this case. Let v_1 denote the largest threshold of $Q_{\{i_1\}\cup\mathcal{I}}$, which is smaller than $v_{i_2}^{j_2-1}$. Consequently, $v_{i_1}^{j_1-1} \leq v_1 < v_{i_2}^{j_2-1}$. Also let $v_2 = \min\{v, v_{i_1}^{j_1}\}$ (in

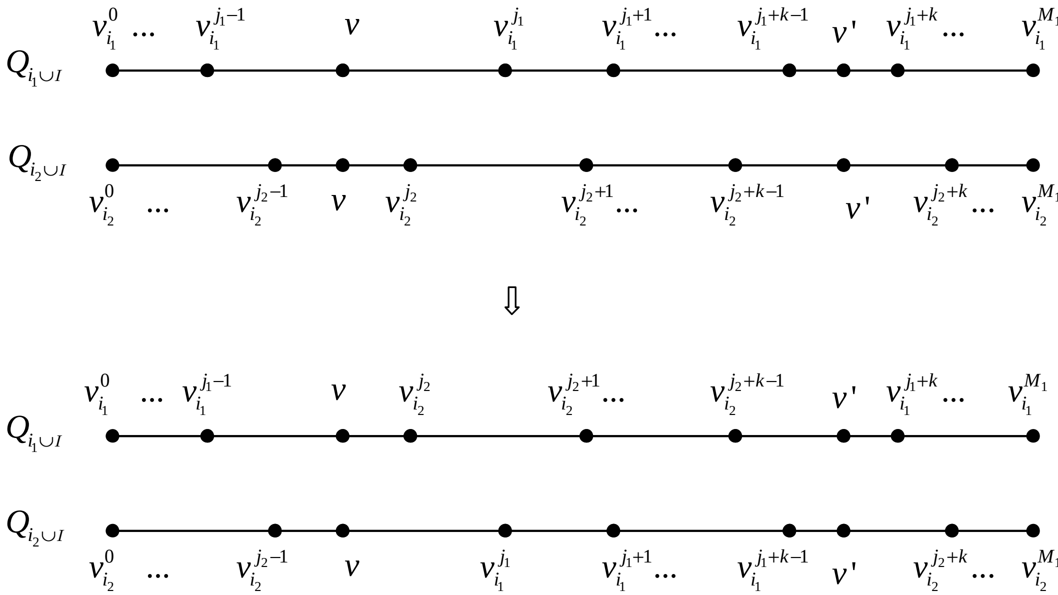

Fig. 4. The effect of thresholds' interchange on partitions $Q_{\{i_1\}\cup\mathcal{I}}$ and $Q_{\{i_2\}\cup\mathcal{I}}$ in **Case 1**.

Figure 5, $v_2 = v$). As an effect of the threshold interchange, the old cells of $Q_{\{i_2\}\cup\mathcal{I}}$ situated between $v_{i_2}^{j_2}$ and v' are exchanged with the old cells of $Q_{\{i_1\}\cup\mathcal{I}}$ situated between v_2 and v'. This exchange does not affect the expected distortion. Additionally, the old cell $(v_{i_2}^{j_2-1}, v_{i_2}^{j_2}]$ of $Q_{\{i_2\}\cup\mathcal{I}}$, is transformed into $(v_{i_2}^{j_2-1}, v_2]$, and the old cell $(v_1, v_2]$ of $Q_{\{i_1\}\cup\mathcal{I}}$, is transformed into $(v_1, v_{i_2}^{j_2}]$. No other modifications occur. Let Δ denote the change in expected distortion due to the modifications in $Q_{\{i_2\}\cup\mathcal{I}}$ and $Q_{\{i_1\}\cup\mathcal{I}}$. Then

$$\Delta = \omega_{\{i_1\}\cup\mathcal{I}}[D(v_{i_2}^{j_2-1}, v_2) + D(v_1, v_{i_2}^{j_2}) - D(v_{i_2}^{j_2-1}, v_{i_2}^{j_2}) - D(v_1, v_2)]. \tag{35}$$

Because $v_1 < v_{i_2}^{j_2-1} < v_{i_2}^{j_2} < v_2$, by applying Lemma 3 and using the fact that $\omega_{\{i_1\}\cup\mathcal{I}} \geq 0$, it follows that $\Delta \leq 0$.

Case 3 ($v < v_{i_2}^{j_2}$ and $v_{i_1}^{j_1+k-1} > v'$), **Case 4** ($v > v_{i_2}^{j_2}$ and $v_{i_1}^{j_1+k-1} > v'$) and the case when $\mathcal{V}$ is empty can be treated by similar arguments. Note that equalities $v = v_{i_2}^{j_2}$ and $v_{i_1}^{j_1+k-1} = v'$ can never hold due to Theorem 3.

In conclusion the threshold interchange does not increase the expected distortion of the K-DSQ. Moreover, if the error function $f(\cdot)$ is additionally continuously differentiable, the expected distortion strictly decreases. With these, the proof is completed. $\square$

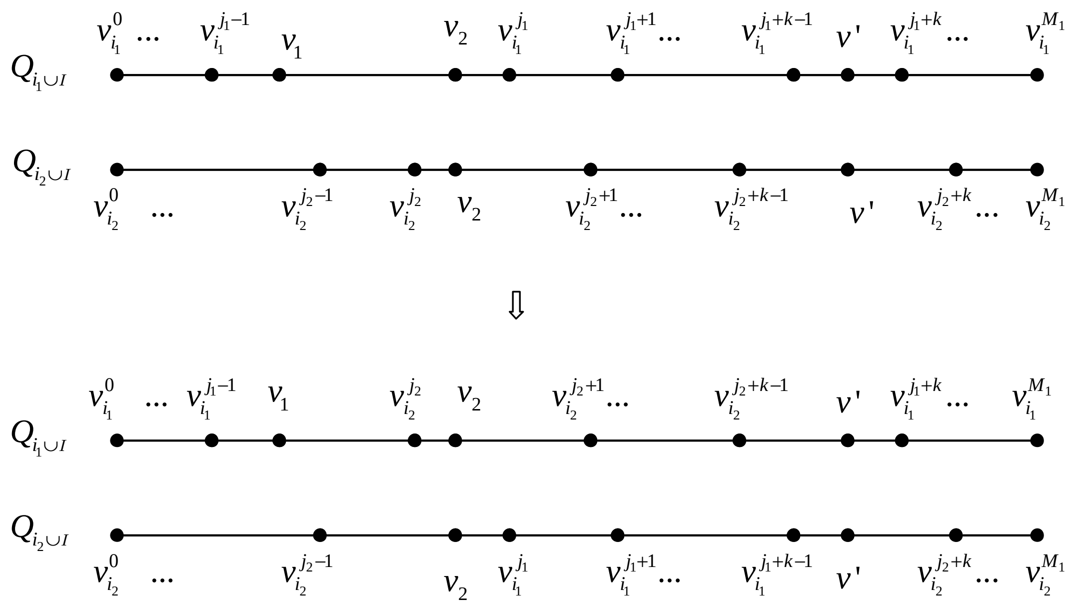

Fig. 5. The effect of thresholds' interchange on partitions $Q_{\{i_1\}\cup\mathcal{I}}$ and $Q_{\{i_2\}\cup\mathcal{I}}$ in **Case 2**.

Relations (27) define the following IA: $h:\{1,2,\cdots,K(M_1-1)+1\}\to\{1,2,\cdots,M_1\}^K$ with

$$h(l)=(\underbrace{j_l+1,\cdots,j_l+1}_{i_l},\underbrace{j_l,\cdots,j_l}_{K-i_l}) \tag{36}$$

where $j_l=\lfloor(l-1)/K\rfloor+1$ and $i_l=l-1-(j_l-1)K$, for all $l, 1\le l\le K(M_1-1)$, and $h(K(M_1-1)+1)=(M_1,M_1,\cdots,M_1)$. According to Theorem 4 this IA is optimal for symmetric convex K-DSQ. Moreover, when the error function $f(\cdot)$ is additionally continuously differentiable, this IA is the unique optimal IA up to a permutation of subscripts of side quantizers. $\square$

VII. Cell Convexity

It has long been known that fixed-rate single description scalar quantizers can be made optimal with convex cells [13]. However, it was recently shown by Gyorgy and Linder [15] that there exist discrete distributions and an interval of rates, for which the optimal entropy-constrained scalar quantizer cannot have convex cells. On the other hand, the same work proves that such an example cannot be found when the source distribution is continuous, the number of quantizer cells is finite and the error function $f(\cdot)$ is non-decreasing and convex. It is also pointed out in [15] that even for discrete distributions, all the points on the operational R-D curve, which are on its convex hull can be achieved by a convex quantizer.

The non-optimality of convex quantizers is not quite as pathological for K-DSQ as for the single description counterpart. Effros and Muresan [10] proved that for both fixed-rate and entropy-constrained situations there are discrete distributions and weights $\omega_{\mathcal{I}}$ such that the optimal K-DSQ cannot be convex. Such examples can be constructed without too much effort even for continuous distributions. For instance, for the uniform distribution over the interval $[0,1]$ and squared error distortion, the optimal fixed-rate symmetric 2-DSQ whose side quantizers have 2 cells each, cannot be convex if $\omega_1/\omega_{\{1,2\}} < 7/81$. Indeed, let $\mathbf{Q}$ be a convex 2-DSQ with 2 cells in each side quantizer. This 2-DSQ has at most 3 cells in the central partition. By bounding from below the distortion of each component quantizer by the lowest distortion achievable at the corresponding rate, we obtain that $\bar{D}(\mathbf{Q}) > \frac{1}{24}\omega_1 + \frac{1}{108}\omega_{\{1,2\}}$. Let now $\mathbf{Q}'$ be the non-convex 2-DSQ whose side partitions are Q_1': $[0,1/2],(1/2,1]$, and Q_2': $[0,1/4]\cup(1/2,3/4],(1/4,1/2]\cup(3/4,1]$. Assume that $\mathbf{Q}'$ has optimal decoder. Then $\bar{D}(\mathbf{Q}') = \frac{17}{192}\omega_1 + \frac{1}{192}\omega_{\{1,2\}}$. Clearly, when $\omega_1/\omega_{\{1,2\}} < 7/81$, we have $\bar{D}(\mathbf{Q}) > \bar{D}(\mathbf{Q}')$.

However, the above results do not rule out the possibility that for many practically important distributions, weights $\omega_{\mathcal{I}}$ and rate constraints, the optimal K-DSQ may have convex cells. A very simple example is the case of fixed-rate MRSQ for a uniform source. Also, it was shown by Vaishampayan [26] (for fixed-rate symmetric 2-DSQ) and by Effros and Muresan [10] (for general case of K-DSQ) that for the squared distance distortion measure, convexity of cells in the central partition does not prevent the K-DSQ from being optimal.

Intuitively, the optimal K-DSQ should be convex when the emphasis in the optimization is on minimizing the side distortions rather than the distortions of other components. This may happen when the ratios $\frac{\omega_i}{\omega_{\mathcal{I}}}$ for $|\mathcal{I}| \geq 2$ are large enough. We conjecture that for any continuous probability distribution $p(x)$, and any K-tuple of positive integers $M_1, \cdots, M_K$ (M_k is the number of cells of side quantizer k), there are finite values $\lambda_{i,\mathcal{I}}$ such that the convexity of the optimal fixed-rate K-DSQ is necessary when $\frac{\omega_i}{\omega_{\mathcal{I}}} > \lambda_{i,\mathcal{I}}$ for all i and $\mathcal{I}$, with $\omega_i, \omega_{\mathcal{I}} \neq 0$. A proof of this statement was given in [6] for fixed-rate symmetric 2-DSQ under the high-resolution assumption ($R \to \infty$) and r-th power distortion measure ($d(x,y) = \mid x-y \mid^r$), when the pdf has a compact support. Precisely, it was shown that under the above conditions, when $\frac{\omega_1}{\omega_{\{1,2\}}} \geq 1/2^{r+1}$, there is an optimal 2-DSQ with all cells convex. This result was obtained by comparison with the high resolution performance of a class of non-convex 2-

DSQ's provided in [28]. As a consequence, in the case when the 2-DSQ is designed for communication over two independent channels, the convex-cell condition does not preclude optimality when the channel probability of success q is at most $\frac{2^{r+1}}{2^{r+1}+1}$, asymptotically in R. Table 1 lists the value of this maximum bound for several values of r. For $r = 2$ the cell convexity will not preclude optimality if the channel has a failure rate of 12% or higher. The larger the value of r, the more relaxed the condition for the side quantizers of optimal 2-DSQ to be convex.

r	$\min \frac{\omega_1}{\omega_{\{1,2\}}}$	$\max q$
1	0.25	0.800
2	0.125	0.888
3	0.0625	0.941
4	0.03125	0.969

TABLE I

MINIMUM VALUE OF THE RATIO OF WEIGHTS $\frac{\omega_1}{\omega_{\{1,2\}}}$ AND MAXIMUM VALUE OF CHANNEL PROBABILITY OF SUCCESS FOR WHICH THE OPTIMAL FIXED-RATE SYMMETRIC 2-DSQ MUST BE CONVEX, IN THE CASE OF CONTINUOUS DISTRIBUTION AND r-TH POWER DISTORTION MEASURE.

For the case of fixed-rate MRSQ, a simple argument shows that at high rates cells convexity does not preclude optimality for any values of the weights $\omega_{\mathcal{I}}$, for the r^{th} power distortion. The argument is based on the analysis of optimal quantization at high rates using the companding approach [2], [1], [3], [20], [14]. As Bennett [2] pointed out, any convex scalar quantizer can be implemented as a compandor. Consider now the optimal companding function (which minimizes the distortion as the rate goes to ∞). Based on this companding function construct K fixed-rate convex quantizers of rates R_1, $R_1 + R_2, \cdots, R_1 + R_2 + \cdots + R_K$, respectively, where $R_i = \log_2 M_i$, for $1 \le i \le K$. These quantizers are embedded, hence they are the active components of a fixed-rate convex MRSQ of K refinement stages. When $R_1 + \cdots + R_i \to \infty$ for all $i, 1 \le i \le K$, the distortion at each stage will become arbitrarily close to the optimal distortion at the corresponding rate. Consequently, the overall expected distortion of the MRSQ will approach the minimal expected distortion, for any values of the weights $\omega_{\mathcal{I}}$.

Theorem 5 states formally the above result. In order to proceed to the statement of the theorem we introduce first some notations. Let $Q_{opt}(R)$ denote the optimal fixed-rate quantizer of rate R, for any

$R > 0$. Moreover, denote

$$\mathcal{J} = \frac{1}{2^r(r+1)}\Big(\int_V^W p^{1/(r+1)}(x)dx\Big)^{r+1}.$$

Then, by [14, Theorem 6.2] we have

$$\lim_{R\to\infty} 2^{rR} D(Q_{opt}(R)) = \mathcal{J}. \tag{37}$$

Theorem 5. Assume that the pdf $p(x)$ is continuous and positive on $[V, W] \cap \mathbf{R}$, and that $p(x) = 0$ outside $[V, W]$. Consider the r-th power distortion function, i.e. $d(x, y) = |x - y|^r$. Moreover, assume that the inequality $\int_V^W |x|^{r+\epsilon} p(x)dx < \infty$ holds for some $\epsilon > 0$ and that there is some $\tau > 0$ such that $p(x)sgn(x)$ is non-increasing in x on each of the intervals $(-\infty, -\tau]$ and $[\tau, \infty)$. Consider an arbitrary sequence of K-tuples of rates (i.e. positive values) $(R_1^{(n)}, R_2^{(n)}, \cdots, R_K^{(n)})_{n\geq 1}$ such that

$$\lim_{n\to\infty}(R_1^{(n)} + R_2^{(n)} + \cdots + R_i^{(n)}) = \infty \text{ for all } 1 \leq i \leq K. \tag{38}$$

Also assume that $2^{R_1^{(n)}+R_2^{(n)}+\cdots+R_i^{(n)}}$ is an integer for any $1 \leq i \leq K$, $n \geq 1$. Let $\mathbf{Q}^{(opt,n)}$ denote the optimal fixed-rate MRSQ achieving the rates $R_1^{(n)}, \cdots, R_K^{(n)}$, i.e., such that each side quantizer i has rate $R_i^{(n)}$, for $1 \leq i \leq K$, $n \geq 1$. Then the following equalities hold

$$\lim_{n\to\infty} 2^{r(R_1^{(n)}+\cdots+R_i^{(n)})} D(Q_{\{1,\cdots,i\}}^{(opt,n)}) = \lim_{n\to\infty} 2^{r(R_1^{(n)}+\cdots+R_i^{(n)})} D(Q_{opt}(R_1^{(n)} + \cdots + R_i^{(n)})) = \mathcal{J},$$

where $Q_{\{1,\cdots,i\}}^{(opt,n)}$ denotes the active component of $\mathbf{Q}^{(opt,n)}$ obtained by intersecting the first i side quantizers. Furthermore, there are convex fixed-rate MRSQ's $\mathbf{Q}^{(n)}$ achieving the rates $R_1^{(n)}, R_2^{(n)}, \cdots, R_K^{(n)}$, for all $n \geq 1$, such that

$$\lim_{n\to\infty} 2^{r(R_1^{(n)}+\cdots+R_i^{(n)})} D(Q_{\{1,\cdots,i\}}^{(n)}) = \mathcal{J}.$$

Proof. Note first that since the rate of component quantizer $Q_{\{1,\cdots,i\}}^{(opt,n)}$ is $R_1^{(n)} + \cdots + R_i^{(n)}$, it follows that

$$\lim_{n\to\infty} 2^{r(R_1^{(n)}+\cdots+R_i^{(n)})} D(Q_{\{1,\cdots,i\}}^{(opt,n)}) \geq \lim_{n\to\infty} 2^{r(R_1^{(n)}+\cdots+R_i^{(n)})} D(Q_{opt}(R_1^{(n)} + \cdots + R_i^{(n)})), \tag{39}$$

for all $1 \leq i \leq K$. To complete the proof we will construct the fixed-rate convex MRSQ $\mathbf{Q}^{(n)}$ using the companding approach. For this, consider the function $g : \mathbf{R} \to [0, 1]$ defined as $g(x) \triangleq \frac{p^{1/(r+1)}(x)}{\int_V^W p^{1/(r+1)}(x)dx}$ for all $x \in \mathbf{R}$. Clearly, g is continuous and positive on $[V, W] \cap \mathbf{R}$. Moreover, $g(x)sgn(x)$ is non-increasing in x on each of the intervals $(-\infty, -\tau]$ and $[\tau, \infty)$. Define now the function $G : [V, W] \cap \mathbf{R} \to [0, 1]$

as $G(x) \triangleq \int_V^x g(t)dt$ for all $x \in [V, W] \cap \mathbf{R}$. Obviously, G is continuous and differentiable and its derivative G' satisfies $G'(x) = g(x)$ for all $x \in [V, W] \cap \mathbf{R}$. Consequently, G is strictly increasing and invertible. Moreover, $\lim_{x \searrow V} G(x) = 0$ and $\lim_{x \nearrow W} G(x) = 1$. Let $h : \mathcal{D}_h \to [V, W] \cap \mathbf{R}$ be the inverse of G ($h = G^{-1}$), where $\mathcal{D}_h$ denotes the domain of definition of h and it satisfies the relations $(0, 1) \subseteq \mathcal{D}_h \subseteq [0, 1]$. Then, for any $R > 0$ such that 2^R is an integer, the function G defines a fixed-rate convex quantizer $Q(R)$ with $N = 2^R$ cells, via the companding approach, as follows. The partition thresholds of $Q(R)$ are $V = t_0 < t_1 < t_2 < \cdots < t_{N-1} < t_N = W$, where $t_i = h(i/N)$ for all $1 \leq i \leq N - 1$. The reproduction value for each cell $(t_{i-1}, t_i]$ is $y_i = h(\frac{2i-1}{2N})$, $1 \leq i \leq N$. By [20, Theorem 1] we have[1]

$$\lim_{R\to\infty} 2^{rR} D(Q(R)) = \frac{1}{2^r(r+1)} \int_V^W \frac{p(x)}{g^r(x)} dx = \mathcal{J}. \tag{40}$$

Finally, note that the quantizers $Q(R_1^{(n)}), Q(R_1^{(n)} + R_2^{(n)}), \cdots, Q(R_1^{(n)} + R_2^{(n)} + \cdots R_K^{(n)})$ are embedded, hence they are respectively, the active components $Q_1^{(n)}, Q_{\{1,2\}}^{(n)}, \cdots, Q_{\mathcal{K}}^{(n)}$, of a fixed-rate convex MRSQ $\mathbf{Q}^{(n)}$, whose side quantizer i has rate $R_i^{(n)}$, for each $1 \leq i \leq K$, $n \geq 1$. Then (37), (38) and (40) imply that

$$\lim_{n\to\infty} 2^{r(R_1^{(n)}+\cdots R_i^{(n)})} D(Q_{\{1,\cdots,i\}}^{(n)}) = \lim_{n\to\infty} 2^{r(R_1^{(n)}+\cdots R_i^{(n)})} D(Q_{opt}(R_1^{(n)} + \cdots R_i^{(n)})) = \mathcal{J},$$

for all $1 \leq i \leq K$. The above equality together with the optimality of MRSQ $\mathbf{Q}^{(opt,n)}$ and the fact that the weights of its active components are positive imply that relation (39) holds with equality. This observation completes the proof. $\square$

Remark 2. Note that a sufficient condition for (38) to hold is that $\lim_{n\to\infty} R_i^{(n)} = \infty$ for all $1 \leq i \leq K$. However, this is not a necessary condition. For instance, (38) is still valid if $\lim_{n\to\infty} R_1^{(n)} = \infty$ while $\lim_{n\to\infty} R_i^{(n)} = c_i$ with $c_i \in \mathbb{R} \cup \{\infty\}$, for $2 \leq i \leq K$.

Remark 3. As an immediate consequence of the above theorem we obtain the following approximation for the expected distortion of the optimal fixed-rate MRSQ $\mathbf{Q}_{opt}(R_1, R_2, \cdots, R_K)$ of rates $R_1, R_2, \cdots, R_K$, as $R_1 + \cdots + R_i \to \infty$ for all i, when the conditions of Theorem 5 are satisfied,

$$\bar{D}(\mathbf{Q}_{opt}(R_1, R_2, \cdots, R_K)) \approx \mathcal{J} \times \big(\sum_{i=1}^{K} \omega_{\{1,2,\cdots,i\}} 2^{-r(R_1+R_2+\cdots+R_i)}\big).$$

[1]It can be easily checked that all conditions in the hypothesis of Theorem 1 of [20] are satisfied.

Moreover this approximation is achieved by a convex fixed-rate MRSQ.

VIII. Conclusion

Sufficient conditions are proven for global optimality of a locally optimal fixed-rate multiple description scalar quantizer (MDSQ) of convex cells, which are the same as those given by Trushkin [24] for fixed-rate single description scalar quantizer counterpart. This work supports the use of generalized Lloyd-algorithm-type methods for scalar multiple description and multiresolution quantizer (MRSQ) design for log-concave probability density functions, such as generalized Gaussian distributions with shape parameter $\beta \geq 1$.

Moreover we address the problem of optimal index assignment for fixed-rate convex MRSQ and symmetric MDSQ, when cell convexity is assumed. In both cases we prove that at optimality the number of cells in the central partition has to be maximal, as allowed by the side quantizer rates. As long as this condition is fullfilled, any index assignment is optimal for MRSQ, while for symmetric MDSQ, an optimal index assignment is proposed.

The assumption of convex cells is also discussed. Notably, it is proved that cell convexity is asymptotically optimal for MRSQ at high resolution, for the r^{th} power distortion measure.

Appendix

Here we state and prove Lemma 3 which is used in Section 6.

We mention that the first part of the following lemma was proved in [29]. However, we need to repeat its proof in order to make clear the proof of the second part.

Lemma 3. Assume that Conditions A and B hold. Then, for $V \leq x \leq x' < y \leq y' \leq W$ the following inequality holds:

$$D(x,y) + D(x',y') \leq D(x',y) + D(x,y'). \tag{41}$$

Moreover, if f is additionally continuously differentiable and $x < x'$, $y < y'$, then the inequality (41) is strict.

Proof. When $x = x'$ or $y = y'$, relation (41) trivially holds with equality. Assume now that $x < x'$ and

$y < y'$. Let $\mu_1 = \mu(x', y)$ and $\mu_2 = \mu(x, y')$. Then

$$D(x', y) + D(x, y') = \int_{x'}^{y} f(|t - \mu_1|)p(t)dt + \int_{x}^{y'} f(|t - \mu_2|)p(t)dt. \tag{42}$$

Assume first that $\mu_1 \le \mu_2$. We will prove now that

$$\int_{x}^{y} f(|t - \mu_1|)p(t)dt + \int_{x'}^{y'} f(|t - \mu_2|)p(t)dt \le \int_{x'}^{y} f(|t - \mu_1|)p(t)dt + \int_{x}^{y'} f(|t - \mu_2|)p(t)dt. \tag{43}$$

The above relation is equivalent to

$$\int_{x}^{x'} f(|t - \mu_1|)p(t)dt \le \int_{x}^{x'} f(|t - \mu_2|)p(t)dt. \tag{44}$$

Because $x' \le \mu_1 \le \mu_2$ we have $|t - \mu_1| \le |t - \mu_2|$ for all $t \in (x, x')$, and further $f(|t - \mu_1|) \le f(|t - \mu_2|)$ since f is strictly increasing. Because $p(t) \ge 0$ for any t, (44) follows. Clearly, the following inequality also holds

$$D(x, y) + D(x', y') \le \int_{x}^{y} f(|t - \mu_1|)p(t)dt + \int_{x'}^{y'} f(|t - \mu_2|)p(t)dt. \tag{45}$$

Relations (42), (43) and (45) imply inequality (41).

Using the notation $D_{a,b}(\xi) = \int_a^b f(|t - \xi|)p(t)dt$ introduced in Section 3, relation (45) can be written as

$$D_{x,y}(\mu(x, y)) + D_{x',y'}(\mu(x', y')) \le D_{x,y}(\mu_1) + D_{x',y'}(\mu_2). \tag{46}$$

Because $\mu(x, y)$ is the unique value satisfying $D_{x,y}(\mu(x, y)) = \min_{\xi \in [V,W]} D_{x,y}(\xi)$, it follows that, when $\mu(x, y) \ne \mu_1$, we have

$$D_{x,y}(\mu(x, y)) < D_{x,y}(\mu_1),$$

and consequently inequality (46) is strict, which further implies that (41) is strict, too.

For the case when f is additionally continuously differentiable, it was proved in [19] (in the Proof of Lemma 1) that the function $\mu(\cdot, \cdot)$, defined on $[V, W] \times [V, W]$, is strictly increasing in each argument. Thus, since $x < x'$, it follows that $\mu(x, y) < \mu_1$, which, together with the abobe considerations, leads to the strictness of inequality (41).

The case $\mu_1 > \mu_2$ can be treated by similar arguments. $\square$

REFERENCES

[1] V. R. Algazi, "Useful approximations to optimal quantization", *IEEE Trans. Commun. Technol.*, vol. COM-14, pp. 297-301, June 1966.
[2] W. R. Bennett, "Spectra of Quantized Signals", Bell Syst. Tech. J., vol. 27, pp. 446-472, July 1948.
[3] J. A. Bucklew, G. L. Wise, "Multidimensional asymptotic quantization theory with r-th power distortion measures", *IEEE Trans. Inform. Theory*, vol. 28, no. 2, pp. 239-247, Mar. 1982.
[4] H. Brunk and N. Farvardin, "Fixed-rate successively refinable scalar quantizers", *Proc. DCC'96*, Snowbird, Utah, March 1996, pp.250-259.
[5] S. Dumitrescu and X. Wu, "Algorithms for optimal multi-resolution quantization", *J. Algorithms*, 50(2004), pp. 1-22.
[6] S. Dumitrescu, X. Wu, "Lagrangian Optimization of Two-description Scalar Quantizers", *IEEE Trans. Inform. Theory*, vol. 53, no. 11, pp. 3990-4012, Nov. 2008.
[7] S. Dumitrescu and X. Wu, "Optimal two-description scalar quantizer design", *Algorithmica*, vol. 41, no. 4, pp. 300. 269-287, Feb. 2005.
[8] S. Dumitrescu and X. Wu, "On multiple description scalar quantizers with convex cells", *Proc. 9th Canadian Workshop Information Theory*, pp. 215-218, Montreal, June 2005.
[9] M. Effros, D. Dugatkin, "Multiresolution Vector Quantization", *IEEE Trans. Inform. Theory*, vol. 50, no. 12, Dec. 2004.
[10] M. Effros and D. Muresan, "Cell Contiguity in Optimal Fixed-Rate and Entropy-Constrained Network Scalar Quantizers", *Proc. DCC'2002*, pp. 312-321, April 2002.
[11] P. E. Fleischer, "Sufficient Conditions for Achieving Minimum Distortion in A Quantizer", *IEEE Int. Conv. Rec.*, 1964, part 1, pp. 104-111.
[12] M. Fleming, Q. Zhao, and M. Effros, "Network Vector Quantization", *IEEE Trans. Inform. Theory*, vol. 50, no. 8, pp. 1584- 1604, Aug. 2004.
[13] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, 1992.
[14] S. Graf, H. Luschgy, *Foundations of quantization for probability distributions*, Springer-Verlag Berlin Heidelberg, 2000.
[15] A. Gyorgy, and T. Linder, "On the structure of optimal entropy-constrained scalar quantizers", *IEEE Transactions on Information Theory*, vol. 48, pp. 416-427, Feb. 2002.
[16] A. Gyorgy, T. Linder, G. Lugosi,"Tracking the Best Quantizer" *IEEE Trans. Inform. Theory*, vol. 54, no. 4, pp. 1604-1625, April 2008.
[17] H. Jafarkhani, H. Brunk, and N. Farvardin, "Entropy-constrained successively refinable scalar quantization," *Proc. IEEE Data Compression Conference*, pp. 337-346, Mar.1997.
[18] J. C. Kieffer, "Exponential rate of Convergence for Lloyd's Method I", *IEEE Trans. Inform. Theory*, vol. IT-28, no. 2, pp. 205-210, Mar. 1982.
[19] J. C. Kieffer, "Uniqueness of Locally Optimal Quantizer for Log-Concave Density and Convex Error Weighting Function", *IEEE Trans. Inform. Theory*, vol. IT-29, no. 1, pp. 42-47, Jan. 1983.
[20] T. Linder, "On asymptotical optimal companding quantization", *Problems of Control and Information Theory*, vol. 20, no. 6, pp.475-484, 1991.
[21] S. P. Lloyd, "Least squares quantization in PCM", *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 129-137, Mar. 1982.
[22] D. Muresan and M. Effros, "Quantization as histogram segmentation: globally optimal scalar quantizer design in network systems", *Proc. DCC'2002*, pp. 302-311, April 2002.
[23] D. Muresan and M. Effros, "Quantization as histogram segmentation: optimal scalar quantizer design in network systems", *IEEE Trans. Inform. Th.*, vol. 54, no. 1, pp. 344-366, Jan. 2008.
[24] A. V. Trushkin, "Sufficient conditions for uniqueness of a locally optimal quantizer for a class of convex error weighting functions", *IEEE Trans. Inform. Th.*, vol. 28, no. 2, pp. 187-198, March 1982.
[25] A. V. Trushkin, "Monotony of Lloyd's Method II for Log-Concave Density and Convex Error Weighting Function", *IEEE Trans. Inform. Th.*, vol. 30, no. 2, pp. 380-383, March 1984.
[26] V. A. Vaishampayan, "Design of multiple-description scalar quantizers", *IEEE Trans. Inform. Th.*, vol. 39, no. 3, pp. 821-834, May 1993.
[27] V. A. Vaishampayan, J. Domaszewicz, "Design of entropy-constrained multiple-description scalar quantizers", *IEEE Trans. Inform. Th.*, vol. 40, no. 1, pp. 245-250, Jan. 1994.
[28] V. A. Vaishampayan, J.-C. Batllo, "Asymptotic Analysis of Multiple Description Quantizers", *IEEE Trans. Inform. Th.*, vol. 44, no. 1, pp. 278-284, Jan. 1998.
[29] X. Wu and K. Zhang, "Quantizer monotonicities and globally optimal scalar quantizer design", *IEEE Trans. Inform. Theory*, vol. 39, pp. 1049-1053, May 1993.

Sorina Dumitrescu received the B.Sc. degree in 1990, and the Ph.D. degree in 1997, both in mathematics, from University of Bucharest, Romania.

She is currently an Assistant Professor at the Department of Electrical and Computer Engineering, McMaster University, Ontario, Canada. Her research interests include multimedia computing and communications, joint source-channel coding, signal quantization, steganalysis. She currently holds an NSERC University Faculty Award.

Xiaolin Wu received his B.Sc. from Wuhan University, China in 1982, and Ph.D. degree from University of Calgary, Canada in 1988, both in computer science.

He is currently a Professor at the Department of Electrical and Computer Engineering, McMaster University, Ontario, Canada, and holds the NSERC-Dalsa Research Chair in digital cinema. Dr. Wu's research interests include network-aware multimedia communication, joint source-channel coding, signal quantization and compression, and image processing. He has published over one hundred seventy research papers and holds two patents in these fields. He is an associate editor of IEEE Transactions on Multimedia.